This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68, C12P 19/34, C07H 21/02, 21/04 — A1

(11) International Publication Number: **WO 99/51773**

(43) International Publication Date: 14 October 1999 (14.10.99)

(21) International Application Number: PCT/US99/07203

(22) International Filing Date: 31 March 1999 (31.03.99)

(30) Priority Data:
60/080,686 3 April 1998 (03.04.98) US

(71) Applicant: PHYLOS, INC. [US/US]; 128 Spring Street, Lexington, MA 02421 (US).

(72) Inventors: KUIMELIS, Robert, G.; 21 Malbert Road, Brighton, MA 02135 (US). WAGNER, Richard; 1007 Lowell Road, Concord, MA 01742 (US).

(74) Agent: ELBING, Karen, L.; Clark & Elbing LLP, 176 Federal Street, Boston, MA 02110–2214 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: ADDRESSABLE PROTEIN ARRAYS

(57) Abstract

Disclosed herein are arrays of nucleic acid–protein fusions which are immobilized to a solid surface through capture probes which include a non–nucleosidic spacer group and an oligonucleotide sequence to which the fusion (such as an RNA–protein fusion) is bound. Also disclosed herein are solid supports on which these arrays are immobilized as well as methods for their preparation and use (for example, for screening for protein–compound interactions such as protein–therapeutic compound interactions).

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

ADDRESSABLE PROTEIN ARRAYS

Background of the Invention

The invention relates to fixed arrays of nucleic acid-protein fusions and, in particular, RNA-protein fusions, on solid supports.

Certain macromolecules, such as proteins, are known to interact specifically with other molecules based on the three-dimensional shapes and electronic distributions of those molecules. For example, proteins interact selectively with other proteins, with nucleic acids, and with small-molecules. Modern pharmaceutical research relies on the study of these interactions; the development of new drugs depends on the discovery of compounds that bind specifically to biologically important molecules.

The discovery of a single drug candidate can require the screening of thousands of compounds. It is therefore important to be able to screen large numbers of compounds rapidly and efficiently. One method for screening a large number of compounds is to fix possible binding partners, such as proteins, to a solid support.

It is difficult to prepare arrays of isolated proteins on solid supports, however, for a variety of reasons. First of all, proteins cannot always be easily attached to the planar surfaces traditionally used to make other fixed arrays, such as nucleic acid microchips. More importantly, because proteins can interact with the functional groups on the surfaces of these supports, the proximity of the protein to the surface can lead to disruption of the protein structure.

Summary of the Invention

In general, the invention features a solid support including an array of immobilized capture probes; each of the capture probes includes a non-nucleosidic spacer group and an oligonucleotide sequence to which a nucleic acid-protein fusion is bound (for example, hybridized or covalently bound). In preferred embodiments, the nucleic acid-protein fusion is an RNA-protein fusion, and the protein component is encoded by the nucleic acid (for example, the RNA). The spacer group can include a polyalkylene oxide, for example, polyethylene oxide. A preferred spacer group includes hexaethylene oxide. The capture probe may also include a photocleavable linker.

The oligonucleotide sequence can include a modified base, such as 5-propyne pyrimidine. It can also include an internucleotide analog (such as 3'-phosphoramidate) or a carbohydrate modification (such as a 2'-O-methyl group). The nucleic acid-protein fusion can include a hybridization tag sequence. The hybridization tag sequence can also include a modified base, an internucleotide analog, or a carbohydrate modification.

In a preferred embodiment, the capture probe further includes a reactive moiety (for example, a nucleophilic group), such as a primary amino group. In another preferred embodiment, the nucleic acid-protein fusion is covalently linked to the capture probe (for example, by photo-crosslinking); in one preferred approach, this is accomplished by including one or more psoralen moieties in the capture probe or in the capture probe-fusion hybridization reaction mixture. A preferred solid support is a glass or silica-based chip.

In a related aspect, the invention features a solid support including an array of immobilized capture probes; each of the capture probes is attached to the surface of the solid support through a non-nucleosidic spacer group, and

each of the capture probes includes an oligonucleotide sequence to which a nucleic acid-protein fusion (for example, an RNA-protein fusion) is bound (for example, hybridized or covalently bound).

In another related aspect, the invention features a solid support including an array of immobilized capture probes; each of the capture probes includes a non-nucleosidic spacer group and an oligonucleotide sequence to which a ribosome display particle is bound (for example, hybridized or covalently bound).

In yet another related aspect, the invention features a method for preparing a solid support. The method includes the steps of: (a) preparing a capture probe by linking a spacer group to an oligonucleotide sequence; (b) attaching the capture probe to the solid support; and (c) binding (for example, hybridizing or covalently binding) a nucleic acid-protein fusion (for example, an RNA-protein fusion) to the capture probe.

The invention also features a second general method for preparing a solid support. This method includes the steps of: (a) attaching a spacer group to a surface of the solid support; (b) attaching a bifunctional linker to the spacer group; (c) attaching a capture probe to the bifunctional linker; and (d) binding (for example, hybridizing or covalently binding) a nucleic acid-protein fusion (for example, an RNA-protein fusion) to the capture probe.

In a second aspect, the invention features a method for detecting an interaction between a protein and a compound. The method includes the steps of: (a) providing a solid support including an array of immobilized capture probes, where each of the capture probes includes a non-nucleosidic spacer group and an oligonucleotide sequence to which a nucleic acid-protein fusion is bound (for example, hybridized or covalently bound); (b) contacting the solid support with a candidate compound under conditions which allow an

interaction between the protein portion of the nucleic acid-protein fusion and the compound; and (c) analyzing the solid support for the presence of the compound as an indication of an interaction between the protein and the compound.

Alternatively, the invention features another method for detecting an interaction between a protein and a compound; this method involves the steps of: (a) providing a population of nucleic acid-protein fusions; (b) contacting the population of nucleic acid-protein fusions with a candidate compound under conditions which allow an interaction between the protein portion of the nucleic acid-protein fusion and the compound; (c) contacting the product of step (b) with a solid support that includes an array of immobilized capture probes, each of the capture probes including a non-nucleosidic spacer group and an oligonucleotide sequence to which a nucleic acid-protein fusion binds (for example, hybridizes or covalently binds); and (d) analyzing the solid support for the presence of the compound as an indication of an interaction between the protein and the compound.

In a preferred embodiment of each of the above methods, the nucleic acid-protein fusion is an RNA-protein fusion. In another preferred embodiment, the compound is labeled. Compounds that can be screened using these methods include, without limitation, proteins, drugs, therapeutics, enzymes, and nucleic acids.

In a third aspect, the invention features an array (for example, an addressable array) of nucleic acid-protein fusions including at least 10^2 different fusions/cm^2. Preferably, the nucleic acid-protein fusions are RNA-protein fusions, and the array includes at least 10^4 different fusions/cm^2.

In a related aspect, the invention features a method for generating an addressable array of molecules. The method involves: (a) providing a solid

support on which an array of nucleic acid molecules is immobilized; (b) contacting the solid support with a population of addressable molecules; and (c) allowing the addressable molecules to orient themselves on the solid support by sequence-dependent recognition and binding of the immobilized nucleic acid molecules.

In preferred embodiments of this method, the addressable array of molecules is an array of nucleic acid-protein fusions (for example, an array of RNA-protein fusions); the addressable molecules orient themselves on the solid support by base pairing (for example, hybridization) with the immobilized nucleic acid molecules; the solid support is a glass or silica-based chip; and the nucleic acid molecules immobilized on the solid support are capture probes, each including a non-nucleosidic spacer group and an oligonucleotide sequence to which the addressable molecule binds.

As used herein, by an "array" is meant a fixed pattern of immobilized objects on a solid surface or membrane. Typically, the array is made up of nucleic acid-protein fusion molecules (for example, RNA-protein fusion molecules) bound to capture nucleic acid sequences which themselves are immobilized on the solid surface or membrane. The array preferably includes at least 10^2, more preferably at least 10^3, and most preferably at least 10^4 different fusions, and these fusions are preferably arrayed on a 125 x 80 mm, and more preferably on a 10 x 10 mm, surface. By an "addressable array" is meant that the locations, or addresses, on the solid support of the members of the array (for example, the nucleic acid-protein fusions) are known; the members of the array are referred to as "addressable molecules" and are utilized in methods for screening for subsequent molecular interactions (for example, for screening for interactions between the addressable nucleic acid-protein fusions and candidate therapeutics).

By "nucleic acid-protein fusion" is meant a nucleic acid covalently bound to a protein. By "nucleic acid" is meant any two or more covalently bonded nucleotides or nucleotide analogs or derivatives. As used herein, this term includes, without limitation, DNA, RNA, and PNA. By "protein" is meant any two or more amino acids, or amino acid analogs or derivatives, joined by peptide or peptoid bond(s), regardless of length or post-translational modification. As used herein, this term includes, without limitation, proteins, peptides, and polypeptides.

By "hybridization tag" is meant a non-coding oligonucleotide sequence that differs sufficiently in sequence from other nucleic acid sequences in a given population or reaction mixture that significant cross-hybridization does not occur. When multiple hybridization tags are utilized in a single reaction mixture, these tags also preferably differ in sequence from one another such that each has a unique binding partner under the conditions employed.

By a "population" is meant more than one molecule.

By a "solid support" is meant any solid surface including, without limitation, any chip (for example, silica-based, glass, or gold chip), glass slide, membrane, bead, solid particle (for example, agarose, sepharose, or magnetic bead), column (or column material), test tube, or microtiter dish.

Brief Description of the Drawings

Figure 1 is a drawing showing the silylation of a glass surface, the derivatization of the resulting amino groups, and the attachment of a capture probe to the modified surface.

Figure 2 is a drawing illustrating a capture probe containing a non-nucleosidic spacer group and a reactive moiety.

Figure 3 is a schematic diagram of the layout of the FLAG and

HA11 fusion chip capture probes utilized in Figures 4 and 5. In this Figure, t7, tag, au1, au5, flag, ha1, irs, and kt3 represent the capture probes CPt7 (positive control), CPtag (positive control), CPau1 (negative control), CPau5 (negative control), CPflag, CPha11, CPirs (negative control), and CPkt3 (negative control), respectively.

Figure 4 is a phosphorimage demonstrating hybridization of nucleic acid-protein fusions (FLAG and HA11) to capture probes immobilized on a chip.

Figure 5 is a fluorimage demonstrating hybridization of nucleic acid-protein fusions (FLAG and HA11) to capture probes immobilized on a chip and subsequent recognition with anti-HA11 monoclonal antibodies.

Figure 6 is a schematic diagram of the layout of the Myc fusion chip capture probes utilized in Figures 7 and 8. In this Figure, capture probes CP01, CP33, CP80, CP125, CPmm, and CPns (described herein) were arranged on the chip as follows: CP01 at locations A1, B1, C1, A4, B4, and C4; CP33 at locations D1, E1, F1, D4, E4, and F4; CP80 at locations A2, B2, C2, A5, B5, and C5; CP125 at locations D2, E2, F2, D5, E5, and F5; CPmm at locations A3, B3, C3, A6, B6, and C6; and CPns at locations D3, E3, F3, D6, E6, and F6.

Figure 7 is a phosphorimage demonstrating hybridization of nucleic acid-protein fusions (Myc) to capture probes immobilized on a chip.

Figure 8 is a fluorimage demonstrating hybridization of nucleic acid-protein fusions (Myc) to capture probes immobilized on a chip and subsequent recognition with anti-Myc monoclonal antibodies.

Description of the Preferred Embodiments

The invention features support-based, addressable arrays of proteins, and methods for preparing and using these arrays. The arrays are prepared by

fixing oligonucleotide sequences, the capture probes (or capture oligos), to a support in a defined array. The capture probes are then used to bind nucleic acid-protein fusions, such as RNA-protein fusions. Such binding may occur through base pairing (for example, through Watson-Crick base pairing, pseudo Watson-Crick base pairing involving modified bases, or Hoogsteen base pairing) between the nucleic acid component of the fusion and a complementary capture probe, or may occur through any other type of sequence-dependent recognition and binding of the capture probe (including, without limitation, polyamide-mediated nucleic acid groove binding or specific binding by nucleic acid-binding proteins such as transcription factors). The result of the binding interactions between the fusions and the capture probes is a defined, addressable array of proteins attached to a solid support.

A variety of materials can be used as the solid support. Examples of such materials include polymers (e.g., plastics), aminated surfaces, gold coated surfaces, nylon membranes, polyacrylamide pads deposited on solid surfaces, silicon, silicon-glass (e.g., microchips), silicon wafers, and glass (e.g., microscope slides). Microchips, and particularly glass microchips, represent a preferred solid support surface.

If the surface is not already aminated, it can be modified to provide a layer of amino groups. For example, a glass microscope slide can be treated with a silylating agent such as trialkoxyaminosilane to provide a surface of primary amino groups that exists as a monolayer or 3-8 molecular layers. This reaction is illustrated in Figure 1. The silane-treated surface is then derivatized with a homobifunctional or heterobifunctional linker that permits the attachment of oligonucleotides at discrete positions. Phenylene 1,4-diisothiocyanate is a useful homobifunctional linker; amino-surfaces derivatized with this reagent have isothiocyanate functionalities that are

available to covalently react with the primary amino groups on the termini of oligonucleotides to form stable thiourea bonds, as shown in Figure 1.

The capture probes, i.e., the oligonucleotide sequences that are to be attached to the surface, are selected from the reverse-complements of the nucleic acid components of the nucleic acid-protein fusions (the targets). Capture probes preferably have between 5 and 30 nucleotide units, and more preferably have about 20 nucleotide units. Considerations for the selection of the exact sequence for a particular capture probe include melting temperature (Tm), interference from competing target sequences, and potential secondary structure in the target sequence. Ideally, each unique capture probe has the same Tm, i.e., they are isoenergetic, so a single hybridization and washing temperature can be used successfully for all capture-target pairs. Commercially available computer programs (e.g., Oligo 4.0) can be used to help identify sets of capture probes with similar thermodynamic properties based on nearest neighbor treatments.

The capture probes are modified before they are attached to the surface. One or more non-nucleosidic spacers, such as polyethylene oxide, are added to the terminus of the oligo. Preferably, 1-20 spacers and, most preferably, 4 spacers are utilized. These spacers may be added to either the 5' or preferably the 3' end of the oligonucleotide. A nucleophilic moiety is then attached to the spacer group. The result is a derivatized capture probe, as shown in Figure 2. A preferred spacer monomer includes hexaethylene oxide.

Non-nucleosidic spacers are preferred over nucleosidic spacers, such as poly-T, because non-nucleosidic spacers have greater flexibility. In addition, their physical properties can be tailored relatively easily, and it is possible to minimize specific and non-specific nucleic acid interactions.

The spacers provide physical separation between the oligonucleotide

and the solid surface and prevent interaction of the proteins with the support surface. This separation is important to ensure effective hybridization between the support-bound capture probe and the nucleic acid-protein fusion. In addition, the separation helps to minimize denaturation of the protein; the proteins are therefore able to adopt their native folded structures and remain functional.

Alternatively, the spacer groups can be attached directly to the solid support surface, instead of to the capture probes. For example, the spacer group can be attached to the amino groups on the surface. The bifunctional linker can then be attached to the other end of the spacer group.

In addition to spacer groups, the capture probes may contain modifications that improve their hybridization properties and mismatch discrimination. For example, they may contain base analogs, such as 5-propyne pyrimidines, internucleotide analogues such as peptide nucleic acids (PNA), in which the bases are connected by peptide-like linkages, or carbohydrate modifications.

The capture probes are suspended in an aqueous alkaline solution, then applied to defined positions of the support surface; the nucleophilic moieties at the termini of the capture probes react with the active sites of the bifunctional linkers to form covalent bonds. The density of the capture probes can be controlled by adjusting reaction time and oligo concentration. Alternatively, the density can be controlled by doping the solution with capture oligos that lack nucleophilic moieties or doping with simple organic compounds that possess amine functional groups.

The capture probes can be applied using liquid deposition techniques, such as inkjet delivery, robotic quill delivery or spotting, and other similar deposition methods. They can also be applied using manual methods,

such as pipetting. The feature sizes of the capture probes can range from one square micron (e.g., when robotic techniques are used) to one square millimeter (e.g., when a 0.2 microliter pipette is used). The result of the application of the capture probes is a defined, regular array of nucleic acid sequences.

After a sufficient reaction time, the excess capture probe is washed away, and the remaining unreacted isothiocyanate groups are blocked off. Dilute ammonia can be used as the blocking agent, resulting in a surface of phenyl thiourea groups. Blocking agents can also be selected to modify the surface energy, i.e., the hydrophobicity of the solid support surface. The hydrophobicity of the solid support surface is important because it affects the background signal level and the extent of unwanted interaction of the protein portion of a nucleic acid-protein fusion with the surface. Examples of blocking agents that modify hydrophobicity are methylamine, amino alcohols, and suitable amino-containing polyethylene oxide moieties.

Non-covalent blocking agents can also be used to further minimize non-specific interactions between the fusion and the solid support (e.g., glass) surface. Examples of such blocking agents include non-specific proteins such as BSA or casein, or similar commercially available blocking reagent formulations marketed for use with membranes.

The capture probes arrayed on the surface of the solid support are then bound (for example, by hybridization) to nucleic acid-protein fusions, such as RNA-protein fusions. A solution containing the mixture of fusions is adjusted to an appropriate salt concentration, applied to the surface, and incubated at a suitable temperature to allow for efficient binding (for example, hybridization) between the capture probe and the target sequence. The solution may also contain surfactants such as TWEEN-20, TRITON X-100, or SDS

(Sigma Chemical Co.) at concentrations of about 0.02% to about 1.0%; it may also include non-specific proteins, such as BSA.

The experimental variables of salt concentration, temperature, and hybridization time are a function of the capture oligo design. A preferred range for the salt concentration is 25 mM to 2 M, with a concentration of about 750 mM being especially preferred. A preferred temperature range is from 5°C to 70°C, with 30°C being especially preferred. Preferred reaction times can be from 1 to 24 hours, with 3 hours being especially preferred. The variables for each experiment are determined empirically by standard methods. The hybridization step can be performed in a simple chamber device that constrains the liquid sample and prevents evaporation.

When RNA-protein fusions are utilized as addressable arrays, the solution may also contain one or more components to suppress nuclease degradation of the RNA moiety. Preferred additions include (a) metal chelators (e.g., EDTA or EGTA) at concentrations of between 1 - 10 mM, (b) placental RNase inhibitor protein (Promega) at concentrations of between 0.1 - 1 Unit/μl; and (c) Anti-RNase protein (Ambion) at concentrations of between 0.1 - 1 Unit/μl. A separate strategy to specifically suppress 5'-exonuclease degradation involves capping the 5'-terminus of the fusion RNA with a binding molecule. The capping strategy may be used in conjunction with one or more of the components listed above. In one particular capping approach, a native or analog (e.g., PNA) nucleic acid sequence conplementary to the 5'-terminus of the fusion RNA is added to generate a stable duplex at the 5'-end. The complementary sequence is preferably between 10 - 50 bases in length, and most preferably abount 20 bases in length. This added nucleic acid sequence may also contain pendant groove-binding, intercalating, or cross-linking moieties. Alternatively, native or analog nucleic acid sequences may be added

that form stable intermolecular hairpin, tetraloop, or pseudoknot secondary structures with the 5'-terminus of the RNA. In the latter case, these nucleic acids are preferably about 20 - 100 bases in length, with about 35 bases being especially preferred.

To the extent possible, the mixture of nucleic acid-protein fusions should be free of un-fused nucleic acids. Un-fused nucleic acids that are complementary to the capture probes will compete with the fusions for binding and will limit the amount of a given protein that can be displayed on the solid support. Preferably, at least 1% of the nucleic acid (for example, the RNA message) is fused to protein.

Unique non-coding regions can be incorporated into the nucleic acid component of the fusion for the specific purpose of being "captured" by the capture probe; these non-coding regions are referred to as "hybridization tag sequences." The hybridization tag sequences may include the same analogue units as are described above for the capture probes. In some cases, both the capture probe and the tag sequences can be modified so they hybridize preferentially with each other, thereby minimizing interference from the coding fusion sequences.

Upon completion of the binding step, unbound nucleic acid-protein fusion is washed away with a buffer that has a higher stringency and a lower salt concentration than that used for the hybridization step. Again, the optimal buffer composition is determined empirically by standard methods. What remains upon completion of washing is an addressable array of proteins on the solid surface, attached via sequence-dependent recognition between the nucleic acid component of the fusion and the surface-bound capture oligo. The position of each protein is defined, because each fusion corresponds to the complementary capture probe.

In addition, if desired, the nucleic acid component of the fusion may be covalently linked to a part of the solid support, the linker, or the capture probe. Such covalently linked fusions provide particularly robust and versatile addressable arrays that may be used, for example, in the screening methods described herein. Covalently linked fusion arrays may be generated by any standard approach. According to one general technique, the fusions are addressed to specific locations on a solid surface via hybridization with corresponding capture probes, and a chemical cross-linking or attachment reaction is triggered to fix the location of the fusions on the solid support. One method to achieve such a covalent link involves functionalizing the DNA capture oligos during chemical synthesis with one or more pendant psoralen moieties, preferably positioned near adenosine bases. After hybridizing the nucleic acid-protein fusion (for example, the RNA-protein fusion) to the support-bound capture oligos, the surface is exposed to long-wavelength UV light (for example, at 350 nm). Light of this wavelength triggers a photoreaction between psoralen and an adjacent thymidine or uridine base in the duplex region, forming a cyclobutane linkage and permanently attaching the fusion to the solid support. Alternatively, psoralen itself (i.e., not linked to a capture probe) may be included in the hybridization solution or in a subsequent separate solution. The psoralen molecule intercalates between bases in double-stranded regions. Upon irradiation with long-wavelength UV light, the intercalated psoralen cross-links with thymidine or uridine bases (intrastrand and interstrand) in a bifunctional mode, forming covalent links between the capture probe and the nucleic acid component of the fusion. Other reactive, cross-linking reagents may also be used in place of psoralen in combination with triggering conditions appropriate for those reagents.

Ordered, addressable arrays of peptide fragments can also be

prepared. To prepare these arrays, the fusion library is generated from short synthetic DNA sequences or fragments of cDNAs or genomic DNAs. In another variation, ribosome display particles, such as those described in Gold et al., WO 93/03172, can be hybridized to the solid support to generate the protein array. Again, these particles are immobilized on the solid support through a hybridization reaction between the capture oligo and the protein-coding RNA.

Use

The addressable protein arrays of the present invention have many uses. For example, a library of proteins can be displayed on a support, such as a microchip. This microchip can then be used to identify previously unknown protein-protein interactions. A probe protein can be detectably labeled, for example, with a radioisotope, chromophore, fluorophore, or chemiluminescent species, then incubated with the microchip. After the excess probe protein is washed away, the chip surface is analyzed for signal from the label. Detection of a signal indicates interaction of the labeled protein with one or more unique members of the protein library. The identity of proteins that are able to bind to the probe protein can then be determined from the location of the spots on the chip that become labeled due to binding of the probe. The same approach can also be used to screen protein libraries for protein-ligand interactions and protein-nucleic acid interactions.

Other methods can be used to detect protein-protein, protein-ligand, or protein-nucleic acid interactions. For example, when the solid surface used to form the protein array is a gold layer, surface plasmon resonance (SPR) can be used to detect mass changes at the surface. When gold surfaces are employed, the reactive moiety on the oligonucleotide capture probe is a thiol group (rather than an amino group) and the gold surface need not be

functionalized to achieve capture probe attachment. Mass spectrometry (especially, Maldi-Tof) can also be used to analyze species bound to unique members of the protein library.

Another application of protein arrays is the rapid determination of proteins that are chemically modified through the action of modifying enzymes such as protein kinases, acyl transferases, and methyl transferases. By incubating the protein array with the enzyme of interest and a radioactively labeled substrate, followed by washing and autoradiography, the location and hence the identity of those proteins that are substrates for the modifying enzyme may be readily determined. Further localization of the modification sites can be achieved using ordered displays of fragments of these proteins.

The protein arrays can also be used to identify the unknown protein targets of therapeutically active compounds. For example, a therapeutic compound may be applied to a protein array derived from cellular RNA. Detection of the captured therapeutic compound, either through its bound label or directly (for example, by mass spectrometry or surface plasmon resonance) reveals the compound's binding partner or partners. In addition, arrays can also be used in the development of protein-based diagnostics. For example, a solid support containing a variety of proteins associated with various illnesses can be prepared. A single patient sample, which might contain one or more proteins whose interactions with the support-bound proteins would be indicative of certain illnesses, can then be contacted with the support. Thus, a single sample can be used to simultaneously detect the presence of several conditions, or to distinguish between conditions. Alternatively, addressable arrays may be used to quantify target molecules in a sample. In one particular example, addressable arrays of single chain antibodies or antibody mimics may be used for quantifying a target protein (or proteins) in a biological sample. The arrays

can also be used in the emerging fields of proteomics and functional genomics.

The specific fusions that are identified as binding specifically to a probe molecule can be removed from the support surface. In one method, the fusion is released by disrupting hybridization with the capture probe. In one particular approach, the specified fusion is physically separated from the rest of the fusions, then treated with a denaturing agent, such as a chemical reagent or heat, to disrupt the base pairing with the capture oligo. The liberated fusion is then recovered from the solution.

Alternatively, the entire capture probe can be detached. During solid support preparation, a light-sensitive linker can be used to attach the capture probe to the solid surface. Following identification of the active fusion, a laser beam of the appropriate wavelength can be used to cleave the linker, thus releasing the desired fusion. Following release from the surface by any of the above methods, the fusion can be specifically recovered and manipulated, for example, using PCR, and further characterized.

There now follow particular examples of the preparation of protein arrays according to the invention. These examples are provided for the purpose of illustrating the invention, and should not be construed as limiting.

Example 1: Silylation of a Glass Surface

Select grade, low-iron content, pre-cleaned 1 x 3 inch glass microscope slides (VWR Scientific) are prepared by heating with 1 M hydrochloric or nitric acid for 30 minutes at 70°C. The slides are then subjected to three 5-minutes washes, using fresh distilled water for each wash. A 1% solution of aminopropyltrimethoxysilane (Gelest, Inc.) in 95% acetone/5% water is prepared and allowed to hydrolyze for at least five minutes. The glass slides are immersed in the hydrolyzed silane solution for

2-20 minutes with gentle agitation. Excess silane is removed by subjecting the slides to ten 5-minute washes, using fresh portions of 95% acetone/5% water for each wash, using gentle agitation. The slides are then cured by heating at 110°C for 20-45 minutes.

Example 2: Derivatization with a Homobifunctional Linker

Silane treated slides from Example 1 are immersed in a freshly prepared 0.2% solution of phenylene 1,4-diisothiocyanate (Aldrich Chemical Co.) in 90% DMF/10% pyridine for two hours, with gentle agitation. The slides are washed sequentially with 90% DMF/10% pyridine, methanol, and acetone. After air drying, the functionalized slides are stored at 0°C in a vacuum desiccator over anhydrous calcium sulfate.

Example 3: Synthesis of Capture Probes

Oligonucleotides are chemically synthesized in the 3'→5' direction by coupling standard phosphoramidite monomers with an automated DNA synthesizer. Typically, 500 angstrom controlled-pore glass supports are used at the 0.2 micromole scale. After the desired probe sequence has been assembled (using A, G, C, and T monomers), hexaethylene oxide phosphoramidite monomer (Glen Research) is added to the 5' terminus. The coupling wait time is extended to 15 minutes by modifying the synthesizer program. Additional hexaethylene oxide monomer units are added in the same way. C-6 Amino phosphoramidite (Glen Research) is then added to the 5' terminus; the coupling wait time is again extended to 15 minutes. The acetic anhydride capping step and the final acidic detritylation step are eliminated. Capture probe sequences are cleaved from the solid support and deprotected with ammonium hydroxide, concentrated to dryness, precipitated in ethanol, and purified by reverse-phase

HPLC using an acetonitrile gradient in triethylammonium acetate buffer.

Example 4: Attachment of Capture Probes

The purified, amine-labeled capture probes from Example 3 are adjusted to a concentration of 500 micromolar in 100 mM sodium carbonate buffer (pH 9.0), and are applied to the derivatized glass surface from Example 2 at defined positions. For manual deposition, aliquots of 0.2 microliter each are applied with a pipetman. The array is incubated at room temperature in a moisture-saturated environment for at least two hours. The attachment reaction is terminated by immersing the glass surface in an aqueous 1% ammonia solution for five minutes with gentle agitation. The glass surface is then subjected to three 5-minute washes, using fresh portions of distilled water for each wash. The array is then soaked in 1 M phosphate buffered saline (PBS) solution for 2 hours at room temperature, then rinsed again for 5 minutes in distilled water.

Example 5: Surface Modification

The ammonia solution from Example 4 is replaced with a 1-5% aqueous solution of a different primary amine-containing molecule. A small amount (10%) of methanol or acetonitrile cosolvent is added, if necessary.

The glass surface is then subjected to three 5-minute washes, using fresh portions of distilled water for each wash. The surface is soaked in 1 M phosphate buffered saline (PBS) solution for 2 hours, then washed again for 5 minutes with distilled water. The glass surface is immersed in a dilute, aqueous solution of a protein-containing blocking solution for several minutes, then subjected to three 5-minute washes, using fresh portions of distilled water for each wash. Finally, the surface is air dried.

Example 6: Fusion Hybridization

50 microliters of a solution containing the RNA-protein fusions and consisting of 25 mM Tris-HCl (pH 8.0) and 100 mM potassium chloride are applied to the glass microchip surface in a chamber that can contain and seal the liquid. The solution is maintained at a specific temperature (determined by the capture oligo design) for at least three hours. Excess, non-hybridized RNA-protein fusions are removed by washing with 25 mM Tris-HCl (pH 8.0) and 50 mM potassium chloride for several minutes at the incubation temperature. The protein chip is subjected to two 15-minute washes, using a buffer that is more stringent and contains a lower salt concentration than the buffer used for the hybridization reaction.

Example 7: Generation of an Exemplary FLAG and HA11 Fusion Chip

Using the techniques essentially as described above, exemplary FLAG and HA11 fusion chips were generated as follows.

For silylation of the glass microchip surface, pre-cleaned 1 x 3 inch glass microscope slides (Goldseal, #3010) were treated with Nanostrip (Cyantek) for 15 minutes, 10% aqueous NaOH at 70°C for 3 minutes, and 1% aqueous HCl for 1 minute, thoroughly rinsing with deionized water after each solution. The slides were then dried in a vacuum desiccator over anhydrous calcium sulfate for several hours. A 1% solution of aminopropytrimethoxysilane (Gelest, Inc.) in 95% acetone / 5% water was prepared and allowed to hydrolyze for 20 minutes. The glass slides were immersed in the hydrolyzed silane solution for 5 minutes with gentle agitation. Excess silane was removed by subjecting the slides to ten 5-minute washes, using fresh portions of 95% acetone / 5% water for each wash, with gentle agitation. The slides were then cured by heating at 110°C for 20 minutes.

To derivatize with a homobifunctional linker, the silane treated slides were immersed in a freshly prepared 0.2% solution of phenylene 1,4-diisothiocyanate (Aldrich Chemical Co.) in 90% DMF / 10% pyridine for two hours, with gentle agitation. The slides were washed sequentially with 90% DMF / 10% pyridine, methanol, and acetone. After air drying, the functionalized slides were stored at 0°C in a vacuum desiccator over anhydrous calcium sulfate.

Capture oligos were then designed and synthesized by standard techniques. In particular, the RNA employed to make the FLAG epitope fusion (17 amino acids total) consisted of 5'-r(UAA UAC GAC UCA CUA UAG GGA CAA UUA CUA UUU ACA AUU ACA AUG GAC UAC AAG GAC GAU GAC GAU AAG GGC GGC UGG UCC CAC CCC CAG UUC GAG AAG) (SEQ ID NO: 1). The RNA employed to make the HA11 epitope fusion (20 amino acids total) consisted of 5'-r(UAA UAC GAC UCA CUA UAG GGA CAA UUA CUA UUU ACA AUU ACA AUG UAC CCC UAC GAC GUG CCC GAC UAC GCC GGC GGC UGG UCC CAC CCC CAG UUC GAG AAG) (SEQ ID NO: 2). In addition, in each case, the following DNA linker, which also contained the essential puromycin moiety at its 3'-end, was ligated to the 3'-terminus of the RNA message:

5'-d(AAAAAAAAAAAAAAAAAAAAAAAAAACC) (SEQ ID NO: 3).

Specific, non-interacting, and thermodynamically isoenergetic sequences along the target RNAs were identified to serve as capture points. The software program HybSimulator v2.2 (Advanced Gene Computing Technology, Inc.) facilitated the identifcation and analysis of potential capture probes. A single specific capture probe for each RNA was ultimately identified (CPflag and CPha11). In addition, two sequences common to each RNA (CPt7, CPtag) were also identified to serve as positive controls. Four non-sense

sequences (CPau1, CPau5, CPirs, CPkt3) were generated as well to serve as negative controls. In total, eight unique sequences were selected. These oligonucleotides were prepared so that they could be attached to the chip surface at either the 3'- or 5'-terminus. Therefore, 16 capture probes were prepared comprising eight unique sequences. The following is a list of these capture probe sequences (5' to 3') (SEQ ID NOS: 4-11):

CPt7:	TGTAAATAGTAATTGTCCC
CPtag:	CTTCTCGAACTGGG
CPau1:	CCTGTAGGTGTCCAT
CPau5:	CAGGTAGAAGTCGGT
CPflag:	CATCGTCCTTGTAGTC
CPha11:	CGTCGTAGGGGTA
CPirs:	CCGCTCCTGATGTA
CPkt3:	TCGGGAGGCATTG.

Oligonucleotide capture probes were chemically synthesized in the 3' to 5' direction by coupling standard phosphoramidite monomers using an automated DNA synthesizer (PE BioSystems Expedite 8909). Typically, 500 angstrom controlled-pore glass supports were used at the 0.2 micromole scale. In the case of 5'-attachment, after the desired probe sequence had been assembled (using A, G, C, and T monomers), four hexaethylene oxide phosphoramidite monomers (Glen Research) were added to the 5'-terminus. The coupling wait time was extended to 15 minutes by modifying the synthesizer program. Additional hexaethylene oxide monomer units were added in the same way. C-6 Amino phosphoramidite (Glen Research) was then added to the 5' terminus; the coupling wait time was again extended to 15

minutes. The acetic anhydride capping step and the final acidic detritylation were eliminated. In the case of 3'-attachment, oligonucleotide synthesis began with a controlled-pore glass support bearing orthoganally protected primary hydroxyl and amino functionalities (Glen Research). Chain elongation began on the hydroxyl group, and the amino group remained protected during oligomer assembly, only being unveiled during the final deprotection. The first four monomers to be added were hexaethylene oxide units, followed by the standard A, G, C, and T monomers. All capture oligo sequences were cleaved from the solid support and deprotected with ammonium hydroxide, concentrated to dryness, precipitated in ethanol, and purified by reverse-phase HPLC using an acetonitrile gradient in triethylammonium acetate buffer. Appropriate fractions from the HPLC were collected, evaporated to dryness in a vacuum centrifuge, and then coevaporated with a portion of water.

To attach the purified, amine-labeled capture oligos, the oligos were adjusted to a concentration of 250 micromolar in 50 mM sodium carbonate buffer (pH 9.0) containing 10% glycerol. The oligos were then robotically applied (MicroGrid, BioRobotics) to the derivatized glass surface described above at defined positions in a 5 x 5 x 16 array pattern (384 spots) within a 20 x 20mm area. The layout of these capture probes is shown schematically in Figure 3. A 16-pin tool was used to transfer the liquid, producing 200 micron features with a pitch of 600 microns. Each sub-grid of 24 spots represented a single capture probe (i.e., 24 duplicate spots). The array was incubated at room temperature in a moisture-saturated environment for 12-18 hours. The attachment reaction was terminated by immersing the glass surface in an aqueous 1% ammonia solution for five minutes with gentle agitation. The glass surface was then subjected to three 5-minute washes, using fresh portions of distilled water for each wash. The array was then soaked in a 10X PBS

(phosphate buffered saline) solution for 2 hours at room temperature, and then rinsed again for 5 minutes in distilled water.

RNA-protein fusions between the peptides containing the FLAG and HA11 epitopes and their corresponding mRNAs were produced as generally described by Szostak et al., WO 98/31700; and Roberts and Szostak, Proc. Natl. Acad. Sci. USA 94:12297-12302, 1997. The polymerase chain reaction using Taq polymerase (Promega) was used to amplify the sequences 5'-TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT ACA ATG GAC TAC AAG GAC GAT GAC GAT AAG GGC GGC TGG TCC CAC CCC CAG TTC GAG AAG (SEQ ID NO: 12) and 5'-TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT ACA ATG TAC CCC TAC GAC GTG CCC GAC TAC GCC GGC GGC TGG TCC CAC CCC CAG TTC GAG AAG (SEQ ID NO: 13) for FLAG and HA11, respectively, using the oligonucleotide primers 5'-TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT (SEQ ID NO: 14) and 5'-AGCGGATGCCTTCTCGAACTGGGGGTGGGA (SEQ ID NO: 15). The resulting PCR products were transcribed in vitro using T7 RNA polymerase (Ambion) to produce an mRNA containing the coding region for the FLAG and HA11 epitopes and the TMV untranslated region. This RNA was ligated to a DNA linker 5'-AAA AAA AAA AAA AAA AAA AAA AAA AAA CC (SEQ ID NO: 3) containing a 5' phosphate and a 3' puromycin by T4 DNA ligase (Promega) in the presence of an 80:20 mixture of the following two DNA splints: 5'-TGCAACGACCAACTTTTTTTTTTAGCGCATGC (SEQ ID NO: 16) and 5'-TGCAACGACCAACTTTTTTTTTTNAGCGCATGC (SEQ ID NO: 17), each containing two biotin moieties at the 5' terminus. The resulting RNA-DNA chimera was purified by binding to Immobilized NeutrAvidin (Pierce), washing to remove unligated material, and eluting by displacement

using the sequence 5'-GCATCCGCTAAAAAAAAAAGTTGGTCGTTGC (SEQ ID NO: 18). Subsequent translations were performed in rabbit reticulocyte lysate (Ambion) according to the manufacturer's instructions except that $MgCl_2$ (150 mM) and KCl (425 mM) were added after 30 minutes to promote the formation of the puromycin-peptide bond. The RNA-peptide fusions were then purified by oligo dT affinity chromatography (Pharmacia), quantitated by scintillation counting of the incorporated vs. added ^{35}S methionine (Amersham), and concentrated to a low volume via membrane filtration (MicroCon).

For hybridization of the fusions to the immobilized capture probes, aliquots of each of the FLAG and HA11 fusions, corresponding to 1.0 picomole each, were combined and adjusted to 5X SSC (saline sodium citrate) + 0.02% Tween-20 in a volume of 20 microliters. The solution was applied to the glass chips described above, coverslips were placed on top, and the slides were placed in a moisture-saturated chamber at room temperature. After 18 hours the coverslips were removed, and the slides were washed sequentially with stirred 500 mL portions of 1X SSC + 0.02% Tween-20, 1X SSC + 0.02% Tween-20, and 1X SSC for 5 minutes each, followed by a brief rinse with 0.2X SSC. After removal of liquid the slides were allowed to briefly air-dry.

To detect hybridization, the FLAG and HA11-fusion chip was exposed to a phosphorimage screen (Molecular Dynamics) for 60 hours by direct contact between the screen and the chip. This allowed identification of the areas that contained hybridized fusions, since the peptides contained a ^{35}S methionine radiolabel which was detectable by the phosphor storage screen. As shown in Figure 4, analysis of the phosphorimage revealed that the fusions had successfully hybridized to their respective capture probes targeting specific areas of the RNA message (i.e., CPflag and CPha11). In addition, the four

non-sense capture probes, which were not compelementary to any region of the FLAG or HA11 RNA, did not give any appreciable signal (i.e., CPau1, CPau5, CPirs, CPkt3). The positive control capture probe CPtag produced the expected signal, but the corresponding positive control capture probe CPt7 did not, likely due to degradation (e.g., exonuclease contamination) of the 5'-region of the targeted RNA. These results demonstrated the feasibility of addressing a mixture of peptides (as fusions) to specific locations on the surface of a chip. Both the 3'-attached capture probes and the 5'-attached capture probes were effective.

A duplicate chip was probed with a monoclonal antibody that recognized the HA11 epitope. All of the following steps were performed at 4°C. Nonspecific sites were first blocked with a solution containing 1X PBS (phosphate buffered saline) + 1% BSA (bovine serum albumin, RNAse free grade, Ambion) + 0.02% Tween-20 for 1 hour under a coverslip. The blocking solution was removed and 50 microliters of HA.11 monoclonal antibody (100:1 dilution, Berkeley Antibody Co.) in 1X PBS + 0.02% Tween-20 was applied to the chip under a coverslip. After 2 hours the coverslip was removed, and the chip was washed with three 50mL portions of 1X PBS + 0.02% Tween-20 for 5 minutes each, with gentle agitation. Excess liquid was removed and then 50 microliters of Cy3-labeled goat anti-mouse IgG (400:1 dilution, Amersham Pharmacia Biotech) in 1X PBS + 0.02% Tween-20 was added under a coverslip. After 1 hour the coverslip was removed, and the chip was washed in three 50mL portions of 1X PBS + 0.02% Tween-20 for 5 minutes each, with gentle agitation. Excess liquid was removed, and the chip was allowed to air-dry at room temperature. The chip was subsequently analyzed at 10 micron pixel resolution with a confocal laser scanner (ScanArray 3000, General Scanning) using preset excitation and emission wavelengths tuned to the Cy3

fluorophore. As shown in Figure 5, the resulting fluorimage was in accord with the phosphorimage and demonstrated that the HA11 peptide, which was covalently linked to its RNA message and fixed to the chip surface, was functional and was available to interact with its binding partner (the HA11 monoclonal antibody). Moreover, although both the FLAG-fusion and the HA11-fusion were presented on the chip surface, the HA11 monoclonal antibody was specific for its own epitope. In addition, the 3'-attachment capture probes generally provided a better signal than the 5'-attachment capture probes. Without being bound to a particular theory, this may reflect the greater accessibility of the epitope when it is oriented away from the chip surface.

Example 8: Generation of an Exemplary Myc Fusion Chip

Using the techniques essentially as described above, an exemplary Myc fusion chip was also generated as follows.

For silylation of the glass surface, select grade, low-iron content, pre-cleaned 25 x 75mm glass microscope slides (VWR Scientific, #48311-950) were used as supplied. A 1% solution of aminopropytrimethoxysilane (Gelest, Inc.) in 95% acetone / 5% water was prepared and allowed to hydrolyze for 20 minutes. The glass slides were immersed in the hydrolyzed silane solution for 5 minutes with gentle agitation. Excess silane was removed by subjecting the slides to ten 5-minute washes, using fresh portions of 95% acetone / 5% water for each wash, with gentle agitation. The slides were then cured by heating at 110°C for 20 minutes.

To derivatize with a homobifunctional linker, the silane treated slides were immersed in a freshly prepared 0.2% solution of phenylene 1,4-diisothiocyanate (Aldrich Chemical Co.) in 90% DMF / 10% pyridine for two hours, with gentle agitation. The slides were washed sequentially with

90% DMF / 10% pyridine, methanol, and acetone. After air drying, the functionalized slides were stored at 0°C in a vacuum desiccator over anhydrous calcium sulfate.

The capture oligos were synthesized based on the Myc sequence. In particular, the RNA employed to make the c-myc fusion (33 amino acids total) consisted of the following sequence:

5'-r(UAAUACGACUCACUAUAGGGACAAUUACUAUUUACAAUUACA AUGGGGACAAUUACUAUUUACAAUUACAAUGGCUGAAGAACAGA AACUGAUCUCUGAAGAAGACCUGCUGCGUAAACGUCGUGAACAGC UGAAACACAAACUGGAACAGCUGCGUAACUCUUGCGCU) (SEQ ID NO: 19). In addition, the following DNA linker, which also contains the essential puromycin moiety, was ligated to the 3'-terminus of the RNA message: 5'-d(AAAAAAAAAAAAAAAAAAAAAAAAAACC) (SEQ ID NO: 3). Three non-overlapping and thermodynamically isoenergetic 20-mer sequences along the RNA were identified to serve as capture points. In addition, dA25 (on the ligated DNA) was selected as a fourth target area. The targeted sequences began at nucleotide positions 1, 33, 80, and 125 (CP01, CP33, CP80 and CP125, respectively). A mismatch sequence, derived from target sequence 33 and containing four internal and adjacent nucleotide mismatches, was also designed (CPmm). A non-sense sequence, corresponding to the reverse-orientation of CP33, was also utilized as a negative control (CPns). The following is a list of the capture probe sequences that were employed (5' to 3') (SEQ ID NOS: 20-25):

CP01: TTGTAAATAGTAATTGTCCC

CP33: AGAGATCAGTTTCTGTTCTT

CP80: AGTTTGTGTTTCAGCTGTTC

CP125: TTTTTTTTTTTTTTTTTTTTTTTTT

Cpmm: AGAGATCTCAATCTGTTCTT

Cpns: TTCTTGTCTTTGACTAGAGA

Oligonucleotide capture probes were chemically synthesized in the 3' to 5' direction by coupling standard phosphoramidite monomers with an automated DNA synthesizer (PE BioSystems Expedite 8909). Typically, 500 angstrom controlled-pore glass supports were used at the 0.2 micromole scale. After the desired probe sequence had been assembled (using A, G, C, and T monomers), hexaethylene oxide phosphoramidite monomer (Glen Research) was added to the 5'-terminus. The coupling wait time was extended to 15 minutes by modifying the synthesizer program. Additional hexaethylene oxide monomer units were added in the same way. C-6 Amino phosphoramidite (Glen Research) was then added to the 5' terminus; the coupling wait time was again extended to 15 minutes. The acetic anhydride capping step and the final acidic detritylation were eliminated. Capture oligo sequences were cleaved from the solid support and deprotected with ammonium hydroxide, concentrated to dryness, precipitated in ethanol, and purified by reverse-phase HPLC using an acetonitrile gradient in triethylammonium acetate buffer. Apppropriate fractions from the HPLC were collected, evaporated to dryness in a vacuum centrifuge, and then coevaporated with a portion of water.

To attach these purified, amine-labeled capture oligos, the oligos were adjusted to a concentration of 500 micromolar in 100 mM sodium carbonate buffer (pH 9.0) and were applied to the derivatized glass surface at defined positions in a 6 x 6 array pattern (36 spots) within a 20 x 20mm area (as shown in Figure 6). CP01 was applied to locations A1, B1, C1 and A4, B4, C4. CP33 was applied to locations D1, E1, F1 and D4, E4, F4. CP80 was

applied to locations A2, B2, C2 and A5, B5, C5. CP125 was applied to locations D2, E2, F2 and D5, E5, F5. Cpmm was applied to locations A3, B3, C3 and A6, B6, C6. Cpns was applied to locations D3, E3, F3 and D6, E6, F6. For manual deposition, aliquots of 0.2 microliter each were applied with a pipetman. The array was incubated at room temperature in a moisture-saturated environment for 12-18 hours. The attachment reaction was terminated by immersing the glass surface in an aqueous 1% ammonia solution for five minutes with gentle agitation. The glass surface was then subjected to three 5-minute washes, using fresh portions of distilled water for each wash. The array was then soaked in a 10X PBS (phosphate buffered saline) solution for 2 hours at room temperature, and then rinsed again for 5 minutes in distilled water.

RNA-protein fusions between a 33 amino acid peptide containing the c-myc epitope and its mRNA were produced as described by Szostak et al., WO 98/31700; and Roberts and Szostak, Proc. Natl. Acad. Sci. USA 94:12297-12302, 1997. The polymerase chain reaction using Taq polymerase (Promega) was used to amplify the sequence 5'-AGC GCA AGA GTT ACG CAG CTG TTC CAG TTT GTG TTT CAG CTG TTC ACG ACG TTT ACG CAG CAG GTC TTC TTC AGA GAT CAG TTT CTG TTC TTC AGC CAT (SEQ ID NO: 26) using oligonucleotide primers 5'-AGC GCA AGA GTT ACG CAG CTG (SEQ ID NO: 27) and 5'-TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT ACA ATG GCT GAA GAA CAG AAA CT (SEQ ID NO: 28). The resulting PCR product was transcribed in vitro using T7 RNA polymerase (Ambion) to produce an mRNA containing the coding region for the c-myc epitope and the TMV untranslated region. This RNA was ligated to a DNA linker 5'-AAA AAA AAA AAA AAA AAA AAA AAA AAA CC (SEQ ID NO: 3) containing a 5' phosphate and a 3' puromycin by T4 DNA

ligase (Promega) in the presence of a DNA splint with the sequence TTT TTT TTT TAG CGC AAG A (SEQ ID NO: 29). The resulting 154mer RNA-DNA chimera was purified by denaturing polyacrylamide gel electrophoresis (6% acrylamide). Translation was performed in rabbit reticulocyte lysate (Ambion) according to the manufacturer's instructions except that KCl (500 mM) was added after 30 minutes to promote the formation of the puromycin-peptide bond. The RNA-peptide fusion was purified by oligo dT affinity chromatography (Pharmacia), quantitated by scintillation counting of the incorporated vs. added ^{35}S methionine (Amersham), and dried to a pellet. 2.5 pmol of the c-myc fusion was produced.

To hybridize to the capture probes, the dry myc-fusion pellet was taken up with 20 microliters of 5X SSC (saline sodium citrate) + 0.02% SDS, mixed, and then briefly centrifuged. The solution was applied to the slide described above, a coverslip was placed on top, and the slide was placed in a moisture-saturated chamber at room temperature. After 18 hours the coverslip was removed, and the slide was washed sequentially with stirred 500 mL portions of 5X SSC + 0.02% SDS, 2.5X SSC + 0.01% SSC, 2.5X SSC, and 1.25X SSC for 5 minutes each. After removal of liquid the slide was allowed to briefly air-dry.

To detect hybridization of the Myc fusions, the glass chip was exposed to a phosphorimage screen (Molecular Dynamics) for four hours by direct contact between the screen and the chip. This allowed identification of the areas that contained hybridized myc-fusion, since the myc peptide contained a ^{35}S methionine radiolabel which was detectable by the phosphor storage screen. As shown in Figure 7, analysis of the phosphorimage revealed that the myc-fusion had successfully hybridized to each of the four capture probes that targeted the myc RNA message and DNA linker sequence. In

addition, the non-sense capture probe, which was not complementary to any region of the myc RNA, did not give any appreciable signal. The capture probe sequence that contained several mismatches produced only a small amount of signal. These results demonstrated that it was possible to address a peptide (as a fusion) to a specific location on the surface of a chip.

After phosphorimage analysis, the same chip was probed with a monoclonal antibody that recognized the c-myc epitope. All of the following steps were performed at 4°C. Nonspecific sites were first blocked with a solution containing 1X PBS (phosphate buffered saline) + 1% BSA (bovine serum albumin, Sigma Chemical Co.) + 0.1 unit per microliter RNAse inhibitor (Ambion) for 1 hour under a coverslip. The blocking solution was removed, and 50 microliters of 9E10 monoclonal antibody in 1X PBS (400:1 dilution, Berkeley Antibody Co.) was applied to the chip under a coverslip. After 1 hour the coverslip was removed, and the chip was washed with three 50mL portions of 1X PBS for 5 minutes each, with gentle agitation. Excess liquid was removed, and then 50 microliters of Cy3-labeled goat anti-mouse IgG in 1X PBS (400:1 dilution, Amersham Pharmacia Biotech) was added under a coverslip. After 1 hour the coverslip was removed, and the chip was washed in three 50mL portions of 1X PBS for 5 minutes each, with gentle agitation. Excess liquid was removed, and the chip was allowed to air-dry at room temperature. The chip was subsequently analyzed at 10 micron pixel resolution with a confocal laser scanner (ScanArray 3000, General Scanning) using preset excitation and emission wavelengths tuned to the Cy3 fluorophore. As shown in Figure 8, the resulting fluorimage was in accord with the phosphorimage and demonstrated that the myc peptide, which was covalently linked to its RNA message and fixed to the chip surface, was functional and was available to interact with its binding partner (the monoclonal antibody).

All publications and patents mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication or patent was specifically and individually indicated to be incorporated by reference.

Other Embodiments

From the foregoing description, it will be apparent that variations and modifications may be made to the invention described herein to adopt it to various usages and conditions. Such embodiments are also within the scope of the following claims.

What is claimed is:

Claims

1. A solid support comprising an array of immobilized capture probes, each of said capture probes comprising a non-nucleosidic spacer group and an oligonucleotide sequence to which a nucleic acid-protein fusion is bound.

2. A solid support comprising an array of immobilized capture probes, wherein each of said capture probes is attached to the surface of said solid support through a non-nucleosidic spacer group, and wherein each of said capture probes comprises an oligonucleotide sequence to which a nucleic acid-protein fusion is bound.

3. A solid support comprising an array of immobilized capture probes, each of said capture probes comprising a non-nucleosidic spacer group and an oligonucleotide sequence to which a ribosome display particle is bound.

4. The solid support of claim 1, 2, or 3, wherein said nucleic acid-protein fusion is an RNA-protein fusion.

5. The solid support of claim 1, 2, or 3, wherein said capture probe is bound to said nucleic acid-protein fusion by base pairing.

6. The solid support of claim 1, 2, or 3, wherein said protein is encoded by said nucleic acid.

7. The solid support of claim 1, 2, or 3, wherein said spacer group

comprises a polyalkylene oxide, polyethylene oxide, or hexaethylene oxide.

8. The solid support of claim 1, 2, or 3, wherein said capture probe comprises a photocleavable linker.

9. The solid support of claim 1, 2, or 3, wherein said oligonucleotide sequence comprises a modified base, an internucleotide analog, or a carbohydrate modification.

10. The solid support of claim 9, wherein said modified base is 5-propyne pyrimidine, said internucleotide analog is a 3'-phosphoramidate linkage, or said carbohydrate modification is a 2'-O-methyl group.

11. The solid support of claim 1, 2, or 3, wherein said nucleic acid-protein fusion comprises a hybridization tag sequence.

12. The solid support of claim 11, wherein said hybridization tag sequence comprises a modified base, an internucleotide analog, or a carbohydrate modification.

13. The solid support of claim 1, 2, or 3, wherein said capture probe further comprises a reactive moiety.

14. The solid support of claim 13, wherein said reactive moiety is a primary amino group.

15. The solid support of claim 1, 2, or 3, wherein said solid support

is a glass or silica-based chip.

16. The solid support of claim 1, 2, or 3, wherein said nucleic acid-protein fusion is covalently linked to said capture probe.

17. The solid support of claim 16, wherein said capture probe comprises one or more psoralen moieties.

18. A method for preparing a solid support, said method comprising the steps of:

(a) preparing a capture probe by linking a spacer group to an oligonucleotide sequence;

(b) attaching said capture probe to said solid support; and

(c) binding a nucleic acid-protein fusion to said capture probe.

19. A method for preparing a solid support, said method comprising the steps of:

(a) attaching a spacer group to a surface of said solid support;

(b) attaching a bifunctional linker to said spacer group;

(c) attaching a capture probe to said bifunctional linker; and

(d) binding a nucleic acid-protein fusion to said capture probe.

20. The method of claim 18 or 19, wherein said nucleic acid-protein fusion is an RNA-protein fusion.

21. A method for detecting an interaction between a protein and a compound, said method comprising the steps of:

(a) providing a solid support comprising an array of immobilized capture probes, each of said capture probes comprising a non-nucleosidic spacer group and an oligonucleotide sequence to which a nucleic acid-protein fusion is bound;

(b) contacting said solid support with a candidate compound under conditions which allow an interaction between the protein portion of said nucleic acid-protein fusion and said compound; and

(c) analyzing said solid support for the presence of said compound as an indication of an interaction between said protein and said compound.

22. A method for detecting an interaction between a protein and a compound, said method comprising the steps of:

(a) providing a population of nucleic acid-protein fusions;

(b) contacting said population of nucleic acid-protein fusions with a candidate compound under conditions which allow an interaction between the protein portion of said nucleic acid-protein fusion and said compound;

(c) contacting the product of step (b) with a solid support comprising an array of immobilized capture probes, each of said capture probes comprising a non-nucleosidic spacer group and an oligonucleotide sequence to which a nucleic acid-protein fusion binds; and

(d) analyzing said solid support for the presence of said compound as an indication of an interaction between said protein and said compound.

23. The method of claim 21 or 22, wherein said nucleic acid-protein fusion is an RNA-protein fusion.

24. The method of claim 21 or 22, wherein said compound is

labeled.

25. The method of claim 21 or 22, wherein said compound is a protein, a therapeutic, an enzyme, or a nucleic acid.

26. An array of nucleic acid-protein fusions, said array comprising at least 10^2 different fusions/cm^2.

27. The array of claim 26, wherein said array comprises at least 10^4 different fusions/cm^2.

28. The array of claim 26, wherein said nucleic acid-protein fusions are RNA-protein fusions.

29. A method for generating an addressable array of molecules, said method comprising:

(a) providing a solid support on which an array of nucleic acid molecules is immobilized;

(b) contacting said solid support with a population of addressable molecules; and

(c) allowing said addressable molecules to orient themselves on said solid support by sequence-dependent recognition and binding of said immobilized nucleic acid molecules.

30. The method of claim 29, wherein said addressable array of molecules is an array of nucleic acid-protein fusions.

31. The method of claim 30, wherein said nucleic acid-protein fusions are RNA-protein fusions.

32. The method of claim 29, wherein said sequence-dependent recognition and binding comprises base pairing.

33. The method of claim 29, wherein said solid support is a glass or silica-based chip.

34. The method of claim 29, wherein said nucleic acid molecules immobilized on said solid support are capture probes, each comprising a non-nucleosidic spacer group and an oligonucleotide sequence to which said addressable molecule binds.

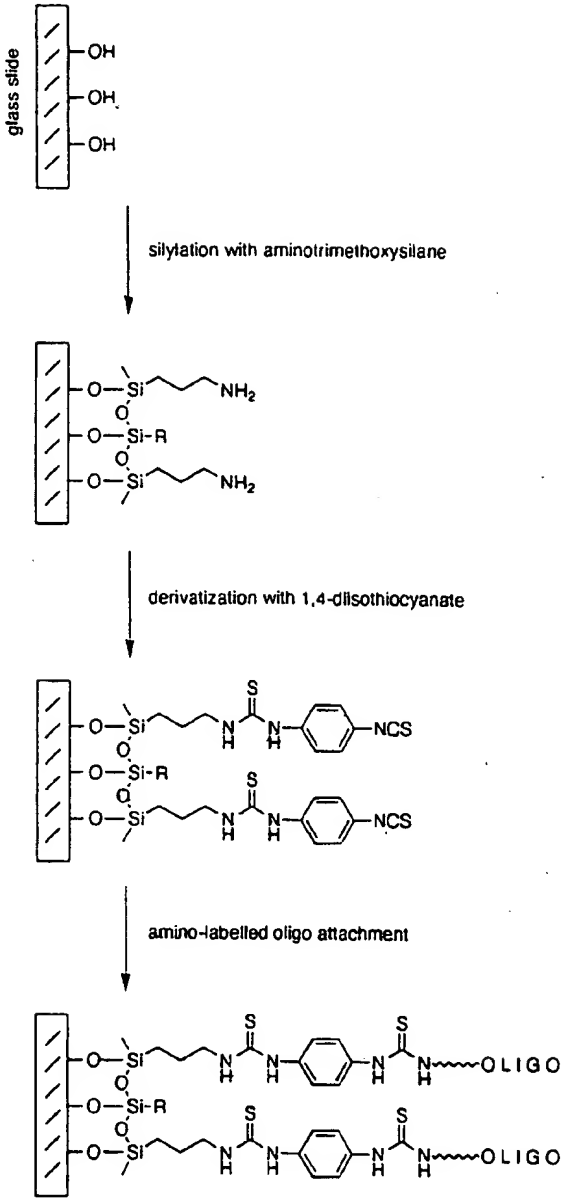

glass slide
OH
OH
OH
silylation with aminotrimethoxysilane
O—Si
NH2
O—Si-R
derivatization with 1,4-diisothiocyanate
S
NCS
amino-labelled oligo attachment
OLIGO

FIGURE 1

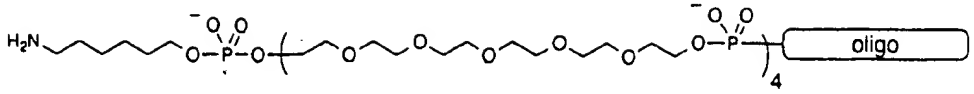


FIGURE 2

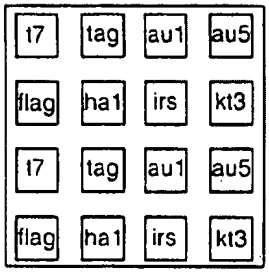


FIGURE 3

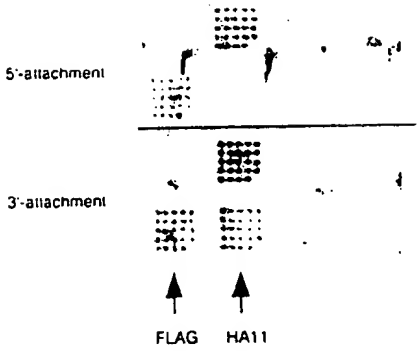


FIGURE 4

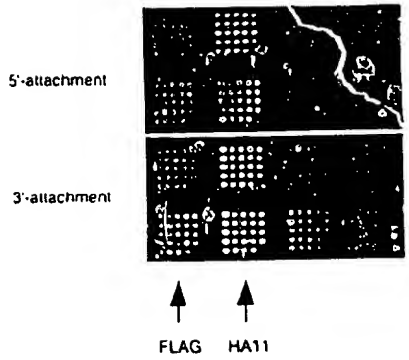


FIGURE 5

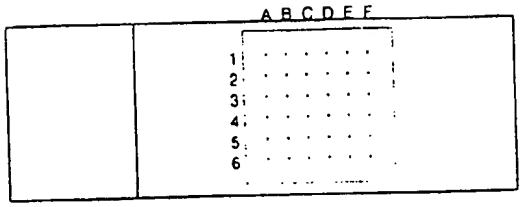


FIGURE 6

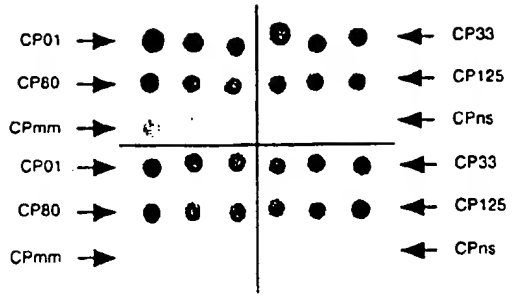


FIGURE 7

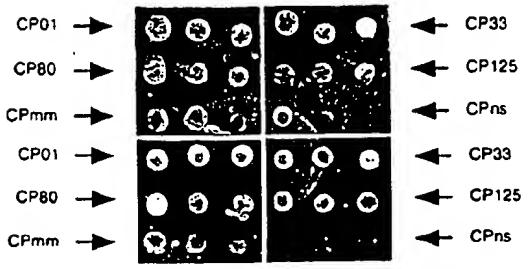


FIGURE 8

SEQUENCE LISTING

<110> Phlyos, Inc.

<120> ADDRESSABLE PROTEIN ARRAYS

<130> 50036/009WO2

<150> 60/080,686
<151> 1998-04-03

<160> 29

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 99
<212> RNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide employed to construct FLAG epitope fusion

<400> 1
uaauacgacu cacuauaggg acaauuacua uuuacaauua caauggacua caaggacgau 60
gacgauaagg gcggcugguc ccacccccag uucgagaag 99

<210> 2
<211> 102
<212> RNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide employed to construct HA11 epitope fusion

<400> 2
uaauacgacu cacuauaggg acaauuacua uuuacaauua caauguaccc cuacgacgug 60
cccgacuacg ccggcggcug gucccacccc caguucgaga ag 102

<210> 3
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for attaching puromycin

<400> 3
aaaaaaaaaa aaaaaaaaaa aaaaaaacc 29

<210> 4

<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for chip attachment

<400> 4
tgtaaatagt aattgtccc 19

<210> 5
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for chip attachment

<400> 5
cttctcgaac tggg 14

<210> 6
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for chip attachment

<400> 6
cctgtaggtg tccat 15

<210> 7
<211> 15
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for chip attachment

<400> 7
caggtagaag tcggt 15

<210> 8
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for chip attachment

<400> 8
catcgtcctt gtagtc 16

<210> 9

<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for chip attachment

<400> 9
cgtcgtaggg gta 13

<210> 10
<211> 14
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for chip attachment

<400> 10
ccgctcctga tgta 14

<210> 11
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for chip attachment

<400> 11
tcgggaggca ttg 13

<210> 12
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> FLAG amplification sequence

<400> 12
taatacgact cactataggg acaattacta tttacaatta caatggacta caaggacgat 60
gacgataagg gcggctggtc ccacccccag ttcgagaag 99

<210> 13
<211> 102
<212> DNA
<213> Artificial Sequence

<220>
<223> HA11 amplification sequence

<400> 13
taatacgact cactataggg acaattacta tttacaatta caatgtaccc ctacgacgtg 60
cccgactacg ccggcggctg gtcccacccc cagttcgaga ag 102

<210> 14
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for PCR

<400> 14
taatacgact cactataggg acaattacta tttacaatt 39

<210> 15
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for PCR

<400> 15
agcggatgcc ttctcgaact gggggtggga 30

<210> 16
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used as a splint containing biotin moiety at 5' terminus

<400> 16
tgcaacgacc aacttttttt tttagcgcat gc 32

<210> 17
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used as a splint containing biotin moiety at 5' terminus

<400> 17
tgcaacgacc aacttttttt ttnagcgcat gc 32

<210> 18
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Oigonucleotide used for elution displacement

<400> 18

gcatccgcta aaaaaaaaag ttggtcgttg c 31

<210> 19
<211> 169
<212> RNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used to make c-myc fusion

<400> 19
uaauacgacu cacuauaggg acaauuacua uuuacaauua caauggggac aauuacuauu 60
uacaauuaca auggcugaag aacagaaacu gaucucugaa gaagaccugc ugcguaaacg 120
ucgugaacag cugaaacaca aacuggaaca gcugcguaac ucuugcgcu 169

<210> 20
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Capture probe sequence

<400> 20
ttgtaaatag taattgtccc 20

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Capture probe sequence

<400> 21
agagatcagt ttctgttctt 20

<210> 22
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Capture probe sequence

<400> 22
agtttgtgtt tcagctgttc 20

<210> 23
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Capture probe sequence

<400> 23
ttttttttttt ttttttttttt ttttt 25

<210> 24
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Capture probe sequence

<400> 24
agagatctca atctgttctt 20

<210> 25
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Capture probe sequence

<400> 25
ttcttgtctt tgactagaga 20

<210> 26
<211> 99
<212> DNA
<213> Artificial Sequence

<220>
<223> c-myc epitope amplification sequence

<400> 26
agcgcaagag ttacgcagct gttccagttt gtgtttcagc tgttcacgac gtttacgcag 60
caggtcttct tcagagatca gtttctgttc ttcagccat 99

<210> 27
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for PCR

<400> 27
agcgcaagag ttacgcagct g 21

<210> 28
<211> 62
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used for PCR

<400> 28
taatacgact cactataggg acaattacta tttacaatta caatggctga agaacagaaa 60
ct 62

<210> 29
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide used as a splint

<400> 29
ttttttttt agcgcaaga 19

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/07203

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12Q 1/68; C12P 19/34; C07H 21/02, 21/04

US CL : 435/6, 91.1,91.2; 536/23.1, 24.3, 24.31,24.32

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/6, 91.1,91.2; 536/23.1, 24.3, 24.31,24.32

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,545,531 A (RAVA et al) 13 August 1996 (13.08.96), see entire document esp. abstract & column 4, lines 4-12.	26,32,33
X	WO 92/10588 A1 (AFFYMAX TECHNOLOGIES N.V..) 25 June 1992 (25.06.92), see entire document.	29,32
X	MASKOS, U. et al. Parallel analysis of oligodeoxyribonucleotide interactions. I. Analysis of factors influencing oligonucleotides duplex formation. Nucleic Acids Research. 1992, Vol. 20, No. 7, pages 1675-1678, see entire document.	29,32
Y	US 5,556,752 A (LOCKHART et al) 17 September 1996 (17.09.96), see entire document.	1-34

[X] Further documents are listed in the continuation of Box C. [ ] See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search
27 JUNE 1999

Date of mailing of the international search report
14 JUL 1999

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
BRADLEY SISSON
Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/07203

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	GUO, Z. et al. Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. Nucleic Acids Research. 1994, Vol. 22, No. 24, pages 5456-5465, see entire document.	1-34
Y	WO 93/03172 A1 (UNIVERSITY RESEARCH CORPORATION) 18 February 1993 (18.02.93), see entire document.	1-34
A,P	WO 98/31700 A1 (THE GENERAL HOSPITAL CORPORATION) 23 July 1998 (23.07.98), see entire document, especially the abstract.	1-34

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US99/07203

B. FIELDS SEARCHED
Electronic data bases consulted (Name of data base and where practicable terms used):

APS, STN, BIOSIS, MEDLINE,CANCERLIT,BIOTECHDS,LIFESCI, CAPLUS, EMBASE
search terms: solid support, chip,microchip,array,probe, oligonucleotide, capture, nucleic acid protein fusion,rna protein fusion, ribosome display particle, spacer, non-nucleosidic spacer, tag

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 March 2001 (22.03.2001)

PCT

(10) International Publication Number
WO 01/20015 A1

(51) International Patent Classification[7]: C12N 15/88, 15/10, C12Q 1/68

(21) International Application Number: PCT/US00/25457

(22) International Filing Date:
18 September 2000 (18.09.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/154,737 17 September 1999 (17.09.1999) US
60/193,580 30 March 2000 (30.03.2000) US

(71) Applicant: WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH [US/US]; Nine Cambridge Center, Cambridge, MA 02142 (US).

(72) Inventor: SABATINI, David, M.; Apartment 6B, 1105 Massachusetts Avenue, Cambridge, MA 02138 (US).

(74) Agents: COLLINS, Anne, J. et al.; Hamilton, Brook, Smith & Reynolds, P.C., Two Militia Drive, Lexington, MA 02421 (US).

(81) Designated States *(national)*: CA, JP.

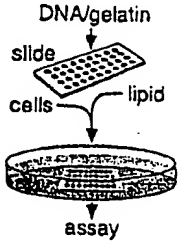
(84) Designated States *(regional)*: European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published:
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: REVERSE TRANSFECTION METHOD

(57) Abstract: A reverse transfection method of introducing DNA of interest into cells and arrays, including microarrays, of reverse transfected cells.



WO 01/20015 A1

REVERSE TRANSFECTION METHOD

BACKGROUND OF THE INVENTION

Genome and expressed sequence tag (EST) projects are rapidly cataloging and cloning the genes of higher organisms, including humans. The emerging challenge is to uncover the functional roles of the genes and to quickly identify gene products with desired properties. The growing collection of gene sequences and cloned cDNAs demands the development of systematic and high-throughput approaches to characterizing the gene products. The uses of DNA microarrays for transcriptional profiling and of yeast two-hybrid arrays for determining protein-protein interactions are recent examples of genomic approaches to the characterization of gene products (Schena, M., *et al., Nature, 10:*623 (2000)). Comparable strategies do not exist to analyze the function, within mammalian cells, of large sets of genes. Currently, *in vivo* gene analysis can be done --on a gene-by-gene scale-- by transfecting cells with a DNA construct that directs the overexpression of the gene product or inhibits its expression or function. The effects on cellular physiology of altering the level of a gene product is then detected using a variety of functional assays.

A variety of DNA transfection methods, such as calcium phosphate coprecipitation, electroporation and cationic liposome-mediated transfection (e.g.,

lipofection) can be used to introduce DNA into cells and are useful in studying gene regulation and function. Additional methods, particularly high throughput assays that can be used to screen large sets of DNAs to identify those encoding products with properties of interest, would be useful to have available.

SUMMARY OF THE INVENTION

Described herein is a strategy for the high throughput analysis of gene function in mammalian cells. A method to create transfected cell microarrays that are suitable for rapidly screening large sets of cDNAs or DNA constructs for those encoding desired products or for causing cellular phenotypes of interest is described. Using a slide printed with sets of cDNAs in expression vectors, a living microarray of cell clusters expressing the gene products has been generated. The cell clusters can be screened for any property detectable on a surface and the identity of the responsible cDNA(s) determined form the coordinates of the cell cluster with a phenotype of interest.

Accordingly, the present invention relates to a method, referred to as a reverse transfection method, in which a defined nucleic acid (a nucleic acid of known sequence or source), also referred to as a nucleic acid of interest or a nucleic acid to be introduced into cells, is introduced into cells in defined areas of a lawn of eukaryotic cells, in which it will be expressed or will itself have an effect on or interact with a cellular component or function. Any suitable nucleic acid such as an oligonucleotide, DNA and RNA can be used in the methods of the present invention. The particular embodiments of the invention are described in terms of DNA. However, it is to be understood that any suitable nucleic acid is encompassed by the present invention.

In one embodiment, the present invention relates to a method in which defined DNA (DNA of known sequence or source), also referred to as DNA of interest or DNA to be introduced into cells, is introduced into cells in defined areas of a lawn of eukaryotic cells, in which it will be expressed or will itself have an effect on or interact with a cellular component or function. In the method, a mixture, defined below, comprising DNA of interest (such as cDNA or genomic DNA incorporated in an expression vector) and a carrier protein is deposited (e.g., spotted

or placed in small defined areas) onto a surface (e.g., a slide or other flat surface, such as the bottoms of wells in a multi-welled plate) in defined, discrete (distinct) locations and allowed to dry, with the result that the DNA-containing mixture is affixed to the surface in defined discrete locations.

Such locations are referred to herein, for convenience, as defined locations. The DNA-containing mixture can be deposited in as many discrete locations as desired. The resulting product is a surface bearing the DNA-containing mixture in defined discrete locations; the identity of the DNA present in each of the discrete locations (spots) is known/defined. Eukaryotic cells, such as mammalian cells (e.g., human, monkey, canine, feline, bovine, or murine cells), bacterial, insect or plant cells, are plated (placed) onto the surface bearing the DNA-containing mixture in sufficient density and under appropriate conditions for introduction/entry of the DNA into the eukaryotic cells and expression of the DNA or its interaction with cellular components. Preferably, the eukaryotic cells (in an appropriate medium) are plated on top of the dried DNA-containing spots at high density (e.g., 1 x $10^5/cm^2$), in order to increase the likelihood that reverse transfection will occur. The DNA present in the DNA-containing mixture affixed to the surface enters eukaryotic cells (reverse transfection occurs) and is expressed in the resulting reverse transfected eukaryotic cells.

In one embodiment of the method, referred to as a "gelatin-DNA" embodiment, the DNA-containing mixture, referred to herein as a gelatin-DNA mixture, comprises DNA (e.g., DNA in an expression vector) and gelatin, which is present in an appropriate solvent, such as water or double deionized water. The mixture is spotted onto a surface, such as a slide, thus producing a surface bearing (having affixed thereto) the gelatin -DNA mixture in defined locations. The resulting product is allowed to dry sufficiently that the spotted gelatin -DNA mixture is affixed to the slide and the spots remain in the locations to which they have become affixed, under the conditions used for subsequent steps in the method. For example, a mixture of DNA in an expression vector and gelatin is spotted onto a slide, such as a glass slide coated with Σ poly-L-lysine (e.g., Sigma, Inc.), for example, by hand or using a microarrayer. The DNA spots can be affixed to the slide by, for example, subjecting the resulting product to drying at room temperature, at

elevated temperatures or in a vacuum-dessicator. The length of time necessary for sufficient drying to occur depends on several factors, such as the quantity of mixture placed on the surface and the temperature and humidity conditions used.

The concentration of DNA present in the mixture will be determined empirically for each use, but will generally be in the range of from about 0.01 μg/μl to about 0.2 μg/μl and, in specific embodiments, is from about 0.02 μg/μl to about 0.10 μg/μl. Alternatively, the concentration of DNA present in the mixture can be from about 0.01 μg/μl to about 0.5 μg/μl , from about 0.01 μg/μl to about 0.4 μg/μl and from about 0.01 μg/μl to about 0.3 μg/μl . Similarly, the concentration of gelatin, or another carrier macromolecule, can be determined empirically for each use, but will generally be in the range of 0.01% to 0.5% and, in specific embodiments, is from about 0.05% to about 0.5%, from about 0.05% to about 0.2% or from about 0.1% to about 0.2%. The final concentration of DNA in the mixture (e.g., DNA in gelatin) will generally be from about 0.02 μg/μl to about 0.1 μg/μl and in a specific embodiment described herein, DNA is diluted in 0.2% gelatin (gelatin in water) to produce a final concentration of DNA equal to approximately 0.05 μg/μl.

If the DNA used is present in a vector, the vector can be of any type, such as a plasmid or viral-based vector, into which DNA of interest (DNA to be expressed in reverse transfected cells) can be introduced and expressed (after reverse transfection) in recipient cells. For example, a CMV-driven expression vector can be used. Commercially available plasmid-based vectors, such as pEGFP (Clontech) or pcDNA3 (Invitrogen), or viral-based vectors can be used. In this embodiment, after drying of the spots containing the gelatin-DNA mixture, the surface bearing the spots is covered with an appropriate amount of a lipid-based transfection reagent and the resulting product is maintained (incubated) under conditions appropriate for complex formation between the DNA in the spots (in the gelatin-DNA mixture) and the lipid-based transfection reagent. In one embodiment, the resulting product is incubated for approximately 20 minutes at 25°C. Subsequently, transfection reagent is removed, producing a surface bearing DNA (DNA in complex with transfection reagent), and cells in an appropriate medium are plated onto the surface. The

resulting product (a surface bearing DNA and plated cells) is maintained under conditions that result in entry of the DNA into plated cells.

A second embodiment of the method is referred to as a "lipid -DNA" embodiment. In this embodiment, a DNA-containing mixture (referred to herein as a lipid-DNA mixture) which comprises DNA (e.g., DNA in an expression vector); a carrier protein (e.g., gelatin); a sugar, such as sucrose; a buffer that facilitates DNA condensation and an appropriate lipid-based transfection reagent is spotted onto a surface, such as a slide, thus producing a surface bearing the lipid-DNA mixture in defined locations. The resulting product is allowed to dry sufficiently that the spotted lipid-DNA mixture is affixed to the slide and the spots remain in the locations to which they have become affixed, under the conditions used for subsequent steps in the method. For example, a lipid-DNA mixture is spotted onto a slide, such as a glass slide coated with Σ poly-L-lysine (e.g., Sigma, Inc.), for example, by hand or using a microarrayer. The DNA spots can be affixed to the slide as described above for the gelatin-DNA method.

The concentration of DNA present in the mixture will be determined empirically for each use, but will generally be in the range of 0.5 μg/μl to 1.0 μg/μl. A range of sucrose concentrations can be present in the mixture, such as from about 0.1M to about 0.4M. Similarly, a range of gelatin concentrations can be present in the mixture, such as from about 0.01% to about 0.05%. In this embodiment, the final concentration of DNA in the mixture will vary and can be determined empirically. In specific embodiments, final DNA concentrations range from about 0.1 μg/μl to about 2.0 μg/μl. If a vector is used in this embodiment, it can be any vector, such as a plasmid, or viral-based vector, into which DNA of interest (DNA to be expressed in reverse transfected cells) can be introduced and expressed (after reverse transfection), such as those described for use in the gelatin-DNA embodiment.

After drying is complete (has occurred to a sufficient extent that the DNA remains affixed to the surface under the conditions used in the subsequent steps of the method), eukaryotic cells into which the DNA is to be reverse transfected are placed on top of the surfaces onto which the DNA-containing mixture has been affixed. Actively growing cells are generally used and are plated, preferably at high

density (such as 1 x $10^5/cm^2$), on top of the surface containing the affixed DNA-containing mixture in an appropriate medium, such as Dulbecco's Modified Eagles Medium (DMEM) containing 10% heat-inactivated fetal serum (IFS) with L-glutamine and penicillin/streptomycin (pen/strep). Other media can be used and their components can be determined based on the type of cells to be transfected. The resulting slides, which contain the dried lipid-DNA mixture and cells into which the DNA is to be reverse transfected, are maintained under conditions appropriate for growth of the cells and entry of DNA, such as an entry of an expression vector containing the DNA, into cells. In the present method, approximately one to two cell cycles are sufficient for reverse transfection to occur, but this will vary with the cell type and conditions used and the appropriate length of time for a specific combination can be determined empirically. After sufficient time has elapsed, slides are assessed for reverse transfection (entry of DNA into cells) and expression of the encoded product or effect of the introduced DNA on reverse-transfected cells, using known methods. This can be done, for example, by detecting immunofluorescence or enzyme immunocytochemistry, autoradiography, in situ hybridization or other means of detecting expression of the DNA or an effect of the encoded product or of the DNA itself on the cells into which it is introduced. If immunofluorescence is used to detect expression of an encoded protein, an antibody that binds the protein and is fluorescently labeled is used (e.g., added to the slide under conditions suitable for binding of the antibody to the protein) and the location (spot or area of the surface) containing the protein is identified by detecting fluorescence. The presence of fluorescence indicates that reverse transfection has occurred and the encoded protein has been expressed in the defined location(s) which show fluorescence. The presence of a signal, detected by the method used, on the slides indicates that reverse transfection of the DNA into cells and expression of the encoded product or an effect of the DNA in recipient cells has occurred in the defined location(s) at which the signal is detected. As described above, the identity of the DNA present at each of the defined locations is known; thus, when expression occurs, the identity of the expressed protein is also known.

Thus, the present invention relates, in one embodiment, to a method of expressing defined DNA, such as cDNA or genomic DNA, in defined locations or

areas of a surface onto which different DNAs, such as DNA in a vector, such as an expression vector, has been affixed, as described herein. Because each area of the surface has been covered/spotted with DNA of known composition, it is a simple matter to identify the expressed protein. In addition, the present method is useful to identify DNAs whose expression alters (enhances or inhibits) a pathway, such as a signaling pathway in a cell or another property of a cell, such as its morphology or pattern of gene expression. The method is particularly useful, for example, as a high-throughput screening method, such as in a microarray format. It can be used in this format for identifying DNAs whose expression changes the phosphorylation state or subcellular location of a protein of interest or the capacity of the cell to bind a reagent, such as a drug or hormone ligand. In a second embodiment, which is also useful as a high-throughput screening method, DNA reverse transfected into cells has an effect on cells or interacts with a cellular component(s) without being expressed, such as through hybridization to cellular nucleic acids or through antisense activity.

Also the subject of this invention are arrays, including microarrays, of defined DNAs spotted onto (affixed to) a surface and array: including microarrays of reverse transfected cells spotted to (affixed to) a surface by the method described herein. Such arrays can be produced by the gelatin-DNA embodiment or the lipid-DNA embodiment of the present method. Arrays of this invention are surfaces, such as slides (e.g., glass or Σ poly-L-lysine coated slides) or wells, having affixed thereto (bearing) in discrete, defined locations DNAs, such as cDNAs or genomic DNA, or cells containing DNA of interest introduced into the cells by the reverse transfection method described herein.

A method of making arrays of the present invention is also the subject of this invention. The method comprises affixing DNAs or reverse transfected cells onto a surface by the steps described herein for the gelatin-DNA embodiment or the lipid-DNA embodiment.

A DNA array of the present invention comprises a surface having affixed thereto, in discrete, defined locations, DNA of known sequence or source by a method described herein. In one embodiment, DNA is affixed to a surface, such as a slide, to produce an array (e.g., a macro-array or a micro-array) by spotting a gelatin-

DNA mixture, as described herein, onto the surface in distinct, defined locations (e.g., by hand or by using an arrayer, such as a micro-arrayer) and allowing the resulting surface bearing the gelatin-DNA mixture to dry sufficiently that the spots remain affixed to the surface under conditions in which the arrays are used. In an alternative embodiment, DNA is affixed to a surface, such as a slide, to produce an array by spotting a lipid-DNA mixture, as described herein, onto the surface in distinct defined locations (e.g., by hand or by using an arrayer, such as a micro-arrayer) and allowing the resulting surface bearing the lipid-DNA mixture to dry sufficiently that the spots remain affixed to the surface under the conditions in which the arrays are used. This result in production of a surface bearing (having affixed thereto) DNA-containing spots.

An array of reverse transfected cells can also be produced by either embodiment described herein. In the gelatin-DNA embodiment, the steps described above for producing DNA arrays are carried out and subsequently, the surface bearing the DNA-containing spots is covered with an appropriate amount of a lipid-based transfection reagent and the resulting product is maintained (incubated) under conditions appropriate for complex formation between DNA in the spots and the reagent. After sufficient time (e.g., about 20 minutes at 25°C) for complex formation to occur, transfection reagent is removed, producing a surface bearing DNA and cells in an appropriate medium are added. The resulting product (a surface bearing DNA and plated cells) is maintained under conditions that result in entry of DNA into plated cells, thus producing an array (a surface bearing an array) of reverse transfected cells that contain defined DNA and are in discrete, defined locations on the array. Such cell arrays are the subject of this invention.

In the lipid-DNA embodiment, the steps described above for producing DNA arrays are carried out and subsequently (after drying is sufficient to affix the DNA-containing spots to the surface, such as a slide or well bottom), cells are plated on top of the surface bearing the DNA-containing spots and the resulting slides, which contain the dried lipid-DNA mixture and cells to be reverse transfected, are maintained under conditions appropriate for growth of the cells and entry of DNA into the cells, thus producing an array (a surface bearing an array) of reverse

transfected cells that contain defined DNA and are in discrete, defined locations on the array. Such arrays are the subject of this invention.

BRIEF DESCRIPTION OF THE DRAWINGS

The file of this patent contains at least one drawing executed in color. Copies of this patent with color drawing(s) will be provided by the Patent and Trademark Office upon request and payment of the necessary fee.

Figure 1 is a schematic representation of one embodiment of the present method of reverse transfection, in which cDNA (HA-GST, HA-FKBP12 or myc-FRB) in an expression vector (prk5) was introduced into cells by the following procedures: combining cDNA in an expression vector, a lipid-based transfection reagent and a carrier protein, to produce a mixture; spotting the mixture onto a glass slide; allowing the spotted mixture to dry on the slide surface; plating human embryonic kidney (HEK 293T) cells into which cDNA is to be introduced onto the slide; maintaining the resulting slide under conditions appropriate for reverse transfection to occur; and detecting immunofluorescence using a fluorescently labeled antibody that binds HA but not myc, demonstrating the presence and location of expressed cDNA.

Figure 2 shows the results of reverse transfection of HEK293T cells with HA-GST, as demonstrated using anti-HA imunofluorescence.

Figure 3 shows the results of reverse transfection of HEK293T cells with pBABE EGFP, as demonstrated by detecting endogenous fluorescence of EGFP.

Figure 4A is a schematic for making transfected cell microarrays using a well-less transfection of plasmid DNAs in defined areas of a lawn of mammalian cells. Plasmid DNA dissolved in an aqueous gelatin solution is printed on a glass slide using a robotic arrayer. The slide is dried and the printed array covered with a lipid transfection reagent. After removal of the lipid, the slide is placed in a culture dish and covered with cells in media. The transfected cell microarray forms in 1-2 days and is then ready for downstream assays. An alternative method in which the lipid is added to the DNA/gelatin solution prior to printing is also described.

Figure 4B is a GFP-expressing microarray made from a slide printed in a 12 x 8 pattern with a GFP expression construct.

Figure 4C is a higher magnification image obtained with fluorescence microscopy of the cell cluster boxed in Figure 4B. Scale bar equals 100 μm.

Figure 4D is a graph of GFP cDNA (picograms) versus mean signal intensity +/- S.D. showing expression levels of clusters in a transfected cell microarray are proportional, over a four-fold range, to the amount of plasmid DNA printed on the slide. Arrays were printed with elements containing the indicated amounts of the GFP construct. Amount of DNA assumes a one nanoliter printing volume. After transfection, the mean +/- S.D. of the fluorescence intensities of the cell clusters were determined. Arrays were prepared as described in Example 3 except that the concentration of the GFP expression plasmid was varied from 0.010-0.050 μg/μl while the total DNA concentration was kept constant at 0.050 μg/μl with empty vector (prk5). Cell clusters were photographed and the signal intensity quantitated with Image Quant (Fuji). The fluorescent image is from a representative experiment.

Figure 4E is a scan image showing that by printing mixtures of two plasmids, cotransfection is possible with transfected cell microarrays. Arrays with elements containing expression constructs for HA-GST, GFP or both were transfected and processed for anti-myc immunofluorescence. For immunofluorescence staining the cells were fixed as described in Example 3, permeabilized in 0.1% Triton X-100 in PBS for 15 minutes at room temperature and probed with primary and secondary antibodies as described. Primary antibodies were used for 1 hour at room temperature at the following concentrations: 1:500 anti-HA ascites (BaBCo), 2 μg/ml anti-myc 9E-10 (Calbiochem), 2 μg/ml anti-V5 (Invitrogen), or 10 μg/ml 4G10 anti-phosphotyrosine (Upstate Biotechnologies). The secondary antibody used was Cy3 μg/ml labeled anti-mouse antibody (Jackson Immunoresearch) at 3.1 μg/ml for 40 minutes at room temperature. Panels labeled Cy3 and GFP show location of clusters expressing HA-GST and GFP, respectively. Merged panel shows superimposition of Cy3 and GFP signals and yellow color indicates co-expression. Scale bar equals 100 μm.

Figure 4F is an enlarged view of boxed area of scan image from Figure 4E.

Figure 5A is a laser scan showing detection of the receptor for FK506. Arrays with elements containing expression constructs for GFP, myc-FKBP12 or

both were printed and transfected with HEK293 cells. 5nM dihydro-FK506 [propyyl-^{3}H] (NEN) was added to the culture media 1 hour prior to fixation and processing for immunofluorescence and autoradiography. Slides were process for anti-myc immunofluorescence, scanned at 5 μm resolution and photographed using a fluorescent microscope, and then exposed to tritium sensitive film (Hyperfilm, Amersham) for 4 days. Autoradiographic emulsion was performed as described by the manufacturer (Amersham). Laser scans show expression pattern of GFP and FKBP12 and superimposition of both (merged). Film autoradiography detects binding of tritiated FK506 to the same array (autorad film).

Figure 5B is a higher magnification image obtained by fluorescent microscopy of an FKBP12-expressing cluster (FKBP12). Emulsion autoradiography detects, with cellular resolution, binding of tritiated FK506 to the same cluster (autorad emulsion).

Figure 5C is a scan showing detected components of tyrosine kinase signaling cascades. 192 V5-epitope-tagged cDNAs in expression vectors were printed in two 8 x 12 subgrids named array 1 and 2. For ease of determining the coordinates of cell clusters within the arrays a border around each array was printed with the GFP expression construct. After transfection, separate slides were processed for anti-V5 or anti-phosphotyrosine immunofluorescence and Cy3 and GFP fluorescence detected. Merged images of array 1 show location of clusters expressing V5-tagged proteins (left panel) and having increased levels of phosphotyrosine (right panel). No DNA was printed in coordinates F10-12.

Figure 5D show two examples of the morphological phenotypes detectable in the transfected cell microarrays described in Figure 5C. Clusters shown are E8 and F7 from array 2.

DETAILED DESCRIPTION OF THE INVENTION

A microarray-based system was developed to analyze the function in mammalian cells of many genes in parallel. Mammalian cells are cultured on a glass slide printed in defined locations with solutions containing different DNAs. Cells growing on the printed areas take up the DNA, creating spots of localized transfection within a lawn of non-transfected cells. By printing sets of

complementary DNAs (cDNAs) cloned in expression vectors, micorarrays which comprise groups of live cells that express a defined cDNA at each location can be made. Transfected cell microarrays can be of broad utility for the high-throughput expression cloning of genes, particularly in areas such as signal transduction and drug discovery. For example, as shown herein, transfected cell microarrays can be used for the unambiguous identification of the receptor for the immunosuppressant FK506 and components of tyrosine kinase pathways.

The present invention relates to a method of introducing defined DNAs into cells at specific discrete, defined locations on a surface by means of a reverse transfection method. That is, the present method makes use of DNAs, of known sequence and/or source, affixed to a surface (DNA spots), such as a slide or well bottom, and growing cells that are plated onto the DNA spots and maintained under conditions appropriate for entry of the DNAs into the cells. The size of the DNA spots and the quantity (density) of the DNA spots affixed to the surface can be adjusted depending on the conditions used in the methods. For example, the DNA spots can be from about 100 μm to about 200 μm in diameter and can be affixed from about 200 μm to about 500 μm apart on the surface. The present method further includes identification or detection of cells into which DNA has been reverse transfected. In one embodiment, DNA introduced into cells is expressed in the cells, either by an expression vector containing the DNA or as a result of integration of reverse transfected DNA into host cell DNA, from which it is expressed. In an alternative embodiment of the present method, DNA introduced into cells is not expressed, but affects cell components and/or function itself. For example, antisense DNA can be introduced into cells by this method and affect cell function. For example, a DNA fragment which is anti-sense to an mRNA encoding a receptor for a drug can be introduced into cells via reverse transfection. The anti-sense DNA will decrease the expression of the drug receptor protein, causing a decrease in drug binding to cells containing the anti-sense DNA. In the method, a mixture comprising DNA of interest (such as cDNA or genomic DNA incorporated in an expression vector) and a carrier protein is deposited (e.g., spotted or placed in small defined areas) onto a surface (e.g., a slide or other flat surface, such as the bottoms of wells in a multi-welled plate) in defined, discrete (distinct) locations and allowed to dry,

with the result that the DNA-containing mixture is affixed to the surface in defined discrete locations.

Detection of effects on recipient cells (cells containing DNA introduced by reverse transfection) can be carried out by a variety of known techniques, such as immunofluorescence, in which a fluorescently labeled antibody that binds a protein of interest (e.g., a protein thought to be encoded by a reverse transfected DNA or a protein whose expression or function is altered through the action of the reverse transfected DNA) is used to determine if the protein is present in cells grown on the DNA spots.

The nucleic acid used in the methods of the present invention can be oligonucleotides, DNA and/or RNA. The nucleic acid of interest introduced by the present method can be nucleic acid from any source, such as nucleic acid obtained from cells in which it occurs in nature, recombinantly produced nucleic acid or chemically synthesized nucleic acid. For example, the nucleic acid can be cDNA or genomic DNA or DNA synthesized to have the nucleotide sequence corresponding to that of naturally-occurring DNA. The nucleic acid can also be a mutated or altered form of nucleic acid (e.g., DNA that differs from a naturally occurring DNA by an alteration, deletion, substitution or addition of at least one nucleic acid residue) or nucleic acid that does not occur in nature. Nucleic acid introduced by the subject method can be present in a vector, such as an expression vector (e.g., a plasmid or viral-based vector), but it need not be. Nucleic acid of interest can be introduced into cells in such a manner that it becomes integrated into genomic DNA and is expressed or remains extrachromosomal (is expressed episomally). The nucleic acid for use in the methods of the present invention can be linear or circular and can be of any size. For example, the nucleic acid can be from about 3 kb to about 10kb, from about 5 kb to about 8 kb and from about 6 kb to 7 kb. Nucleic acid introduced into cells by the method described herein can further comprise nucleic acid (e.g., DNA) that facilitates entry of the nucleic acid into cells or passage into the cell nucleus (nuclear localization elements).

The carrier for use in the methods of the present invention can be, for example, gelatin or an equivalent thereof.

Eukaryotic cells, such as mammalian cells (e.g., human, monkey, canine, feline, bovine, or murine cells), bacterial, insect or plant cells, are plated (placed) onto the surface bearing the DNA-containing mixture in sufficient density and under appropriate conditions for introduction/entry of the DNA into the eukaryotic cells and expression of the DNA or its interaction with cellular components. Preferably, the eukaryotic cells (in an appropriate medium) are plated on top of the dried DNA-containing spots at high density (e.g., 0.5-1 x $10^5/cm^2$), in order to increase the likelihood that reverse transfection will occur. For example, the density of cells can be from about 0.3 x $10^5/cm^2$ to about 3 x $10^5/cm^2$, and in specific embodiments, is from about 0.5 x $10^5/cm^2$ to about 2 x $10^5/cm^2$ and from about 0.5 x $10^5/cm^2$ to about 1 x $10^5/cm^2$. The appropriate conditions for introduction/entry of DNA into cells will vary depending on the quantity of cells used.

Two embodiments of the present method are described in detail herein: a DNA-gelatin method, in which a mixture comprising DNA (e.g., DNA in an expression vector, such as, a plasmid-based or viral-based vector) and a carrier protein (e.g., gelatin) is used and a lipid vector-DNA method, in which a mixture comprising DNA, such as DNA in an expression vector (e.g., a plasmid); a carrier protein (e.g., gelatin); a sugar (e.g., sucrose); DNA condensation buffer; and an appropriate lipid-containing transfection reagent is used. Any suitable gelatin which is non-toxic, hydrated, which can immobilize the nucleic acid mixture onto a surface and which also allows the nucleic acid immobilized on the surface to be introduced over time into cells plated on the surface can be used. For example, the gelatin can be a crude protein gelatin or a more pure protein based gelatin such as fibronectin. In addition, a hydrogel, a sugar based gelatin (polyethylene glycol) or a synthetic or chemical based gelatin such as acrylamide can be used.

In the first embodiment, a mixture comprising two components (DNA such as DNA in an expression vector and a carrier protein) is spotted onto a surface (e.g., a slide) in discrete, defined locations or areas and allowed to dry. One example of this embodiment is described in Example 1. After the carrier (e.g., gelatin)-DNA mixture has dried sufficiently that it is affixed to the surface, transfection reagents (a lipofection mixture) and cells to be reverse transfected are added, preferably sequentially. The transfection mixture can be one made from available components

or can be a commercially available mixture, such as Effectene™ (Qiagen), Fugene™ 6 (Boehringer Mannheim) or Lipofectamine™ (Gibco/BRL-Life Technologies). It is added in an appropriate quantity, which can be determined empirically, taking into consideration the amount of DNA in each defined location. A wax barrier can be drawn around the locations on the surface which contain the vector-DNA mixture, prior to addition of the transfection mixture, in order to retain the mixture or the solution can be kept in place using a cover well. Generally, in this embodiment, the transfection reagent is removed, such as by vacuum suctioning, prior to addition of cells into which DNA is to be reverse transfected. Actively growing cells are plated on top of the locations, producing a surface that bears the DNA-containing mixture in defined locations. The resulting product is maintained under conditions (e.g., temperature and time) which result in entry of DNA in the DNA spots into the growing cells. These conditions will vary according to the types of cells and reagents used and can be determined empirically. Temperature can be, for example, room temperature or 37°C, 25°C or any temperature determined to be appropriate for the cells and reagents.

A variety of methods can be used to detect protein expression in the DNA-containing spots. For example, immunofluorescence can be used to detect a protein. Alternatively, expression of proteins that alter the phosphorylation state or subcellular localization of another protein, proteins that bind with other proteins or with nucleic acids or proteins with enzymatic activity can be detected.

In the second embodiment, one example of which is described in Example 2, a mixture comprising DNA in an expression vector; a carrier protein (e.g., gelatin); a sugar (e.g., sucrose); DNA condensation buffer; and a lipid-based transfection reagent is spotted onto a surface, such as a slide, in discrete, defined locations and allowed to dry. Actively growing cells are plated on top of the DNA-containing locations and the resulting surface is maintained under conditions (e.g., temperature and time) which result in entry of DNA in the DNA spots into the growing cells (reverse transfection). Expression of DNA in cells is detected using known methods, as described above.

Any suitable surface which can be used to affix the DNA containing mixture to its surface can be used. For example, the surface can be glass, polystyrene or plastic. In addition, the surface can be coated with, for example, polylysine.

The present invention also encompasses methods of making arrays which comprise DNA affixed to a surface such that when cells are plated onto the surface bearing the DNA, the DNA can be introduced (is introducible) into the cells (i.e., the DNA can move from the surface into the cells). The present invention also encompasses a DNA array comprising a surface having affixed thereto, in discrete, defined locations, DNA of known sequence or source by a method described herein.

The methods of this invention are useful to identify DNAs of interest (DNAs that are expressed in recipient cells or act upon or interact with recipient cell constituents or function, such as DNAs that encode a protein whose function is desired because of characteristics its expression gives cells in which it is expressed). They can be used in a variety of formats, including macro-arrays and micro-arrays. They permit a DNA array to be converted into a protein or cell array, such as a protein or cell microarray.

The present invention is illustrated by the following examples, which are not intended to be limiting in any way.

Example 1 Reverse Transfection: "Gelatin-DNA" Method

Materials

[DNA]: 1μg/μL (eg., HA-GST pRK5, pBABE CMV GFP)

Gelatin (ICN, cat.# 901771): 0.2% stock in ddH_20, all dilutions made in PBS

0.20% gelatin = 0.5g gelatin + 250mL ddH_2O

Effectene Transfection Kit (Qiagen, cat.# 301425)

Plasmid-DNA: grown in 100mL L-amp overnight from glycerol stock, purified by standard Qiaprep Miniprep or Qiagen Plasmid Purification Maxi protocols

Cell Type: HEK 293T cultured in DMEM/10%IFS with L-glut and pen/strep

Diluting and Spotting DNA

- Dilute DNA in 0.2% gelatin* to give final [DNA]=0.05μg/μL**
- Spot DNA/gelatin mix on Σ poly-L-lysine slides using arrayer

- Allow slides to dry in vacuum-dessicator overnight***

* range of gelatin concentration that worked under the conditions used = 0.05% to 0.5%

** range of DNA concentrations that worked under the conditions used = 0.01 μg/μl to 0.10μg/μl

*** range of drying time = 2 hours to 1 week

Adding Tx. Reagents to Gelatin -DNA Spots

- In eppendorf tube, mix 300μL DNA-condensation buffer (EC Buffer)+ 16μL Enhancer
- Mix by vortexing. Incubate for 5 minutes
- Add 50μL Effectene and mix by pipetting
- Draw a wax circular barrier on slide around spots to apply the transfection reagent
- Add 366μL mix to wax-enclosed region of spots
- Incubate at room temperature for10 to 20 minutes
- Meanwhile, split cells to reverse-transfect
- Vacuum-suction off reagent in hood

Place slides in dish and add cells for reverse transfection

Splitting Cells

- Split actively growing cells to [cell] = 10^7 cells in 25mL
- Plate cells on top of slide(s) in square 100x100x15mm petri dish
- Allow reverse transfection to proceed for 40 hours = approx. 2 cell cycles
- Process slides for immunofluorescence

Example 2 Reverse Transfection: "Lipid - DNA" Method

Materials

[DNA]:1μg/μL (eg., HA-GST pRK5, pBABE CMV GFP)

Gelatin (ICN, cat.# 901771): 0.2% stock in ddH_20, all dilutions made in PBS^-

0.05% gelatin = 250μL 0.2% + 750μL PBS^-

Effectene Transfection Kit (Qiagen, cat.# 301425):

EC Buffer in 0.4M sucrose = 273.6µL 50% sucrose + 726.4µL EC Buffer

Plasmid-DNA: grown in 100mL L-amp overnight from glycerol stock, purified by standard Qiaprep Miniprep or Qiagen Plasmid Purification Maxi protocols

Cell Type: HEK 293T cultured in DMEM/10%IFS with L-glut and pen/strep

Reverse Transfection Protocol with Reduced Volume

- Aliquot 1.6µg DNA in separate eppendorf tubes
- Add 15µL of pre-made *DNA-condensation buffer (EC Buffer) with 0.4M sucrose** to tubes
- Add 1.6µL of Enhancer solution and mix by pipetting several times. Incubate at room temperature for 5 minutes
- Add 5µL of Effectene Transfection Reagent to the DNA-Enhancer mix and mix by pipetting. Incubate at room temperature for 10 minutes
- Add 23.2µL of 0.05% gelatin** to total transfection reagent mix (i.e. 1:1 dilution)
- Spot lipid-DNA on Σ poly-L-lysine slides mix using arrayer
- Allow slides to dry in vacuum-dessicator overnight***

Effectene™ kit (Qiagen) used includes Enhancer solution, which was used according to Qiagen's instructions.

* range of sucrose that worked under the conditions used = 0.1M to 0.4M

** range of gelatin concentration that worked under the conditions used = 0.01% to 0.05%

*** range of drying time = 2 hours to 1 week

Splitting Cells

- Split actively growing cells to [cell] = 10^7 cells in 25mL
- Plate cells on top of slide(s) in square 100x100x15mm petri dish
- Allow reverse transfection to proceed for 40 hours = approx. 2 cell cycles
- Process slides for immunofluorescence

Example 3 Transfected Cells Micorarrays: a genomics approach for the analysis of gene products in mammalian cells

Lipid-DNA Method

I. Gelatin Preparation and DNA Purification

Materials:

Gamma-Amino Propyl Silane (GAPS) slides (Corning catalog #2550),

Purified cDNA,

Gelatin, Type B: 225 Bloom (Sigma, catalog #G-9391),

Methods:

0.2% Gelatin was made by incubation in a 60°C water bath for 15 minutes. The gelatin was cooled slowly to 37°C at which point it was filtered through 0.45μm cellular acetate membrane (CA).

Bacterial clones with DNA plasmids were grown in a 96 Deep-Well Dish for 18 to 24 hours in 1.3mL of terrific broth (TB) shaking at 250rpm at 37°C. The plasmids were miniprepped and optical density (OD) was taken. DNA purity, as indicated by final 280nm/260nm absorbance ratio, was greater than 1.7.

Storage:

For storage purposes, gelatin was kept at 4°C and miniprepped DNA kept at -20°C.

II. Sample Preparation and Array Printing

Materials:

Effectene Transfection Reagent (Qiagen catalog #301425),

Sucrose (Life Technologies),

INTEGRID 100mm x 15mm Tissue Culture Square Petri Dishes (Becton Dickinson: Falcon catalog #35-1012),

Costar 384-well plates (VWR catalog #7402),

Stealth Micro Spotting Pins, (Telechem International, Inc. catalog #SMP4),

PixSys 5500 Robotic Arrayer (Cartesian Technologies, Model AD20A5),

Vacuum Dessicator with Stopcock 250mm, NALGENE (VWR catalog #24987-004),
DRIERITE Anhydrous Calcium Sulfate (VWR catalog #22890-229)
Forceps to hold slides,
Human Embryonic Kidney (HEK) 293T cells,
Tissue Culture hood,
Cover Slips (50mm x 25mm),

Methods:

For each DNA-containing spot, 15µl of pre-made DNA-condensation buffer (Buffer EC) with 0.2M to 0.4M sucrose was added to 0.80µg to 1.60µg DNA in a separate eppendorf tube. Subsequently, 1.5µl of the Enhancer solution was added to the tube and mixed by pipetting. This was let to incubate at room temperature for 5 minutes. 5µl Effectene transfection reagent was added, mixed and let to incubate at room temperature for 10 minutes with the DNA-Enhancer mixture. 1X volume of 0.05% gelatin was added, mixed and the appropriate amount was aliquoted into a 384-well plate for arraying purposes.

The PixSys 5500 Robotic Arrayer was used with Telechem's ArrayIt Stealth Pins (SMP4) with each spot spaced 400µm apart with a 50ms to 500ms delay time of the pin on the slide for each spot. A 55% relative humidity environment was maintained during the arraying. A thorough wash step was implemented between each dip into a DNA sample in the 384-well plate to avoid clogging of the pins that would result in missing spots in the array.

In a tissue culture hood, $10x10^6$ Human Embryonic Kidney (HEK) 293T cells were prepared in 25ml DME media with 10% IFS, pen/strep and glutamine for every 3 slides that were to be processed. After arraying, the slides were simply placed array-side facing up on a sterile 100x100x10mm square dish (3 slides per plate) and the cells were poured gently on the slides while avoiding direct pouring on the arrays themselves. If the number of slides were not a multiple of 3, dummy slides were placed to cover the square dish.

The cells were let to grow on the arrays for approximately 2 cell cycles (~40hours for 293T). Subsequently, the slides were very gently rinsed with PBS- in a Coplin jar, and then fixed in 3.7% paraformaldehyde/4.0% sucrose for 20 minutes in a Coplin jar, and then transferred back to a jar with PBS-.

Storage:

After arraying, slides were stored at room temperature in a vacuum dessicator with anhydrous calcium sulfate pellets. After fixation, slides were kept in PBS- at 4°C until analyses were completed (maximum of 5 days).

III. Methods of Detection

Immunofluorescence

Fluorescence Microscopy

Laser Scanning

Radiolabelling and detection with sensitive film or emulsion

If the expressed proteins to be visualized are fluorescent proteins, they can be viewed and photographed by fluorescent microscopy. For large expression array, slides may be scanned with a laser scanner for data storage. If a fluorescent antibody can detect the expressed proteins, the protocol for immunofluorescence can be followed. If the detection is based on radioactivity, the slides can be fixed as indicated above and radioactivity detected by autoradiography with film or emulsion.

Immunofluorescence:

After fixation, the array area was permeabilized in 0.1% Triton X-100 in PBS- for 15 minutes. After two rinses in PBS-, the slides were blocked for 60 minutes, probed with a primary antibody at 1:200 to 1:500 dilution for 60 minutes, blocked for 20 minutes, probed with a fluorescent secondary antibody at 1:200 dilution for 40 minutes. The slides can be transferred to a Coplin jar in PBS- and visualized under an upright fluorescent microscope. After analyses, the slides can be mounted and stored in the dark at 4°C.

To create these microarrays, distinct and defined areas of a lawn of cells were simultaneously transfected with different plasmid DNAs (Figure 4A). This is accomplished without the use of individual wells to sequester the DNAs. Nanoliter volumes of plasmid DNA in an aqueous gelatin solution are printed on a glass slide. A robotic arrayer (PixSys 5500, Cartesian Technologies) equipped with stealth pins (SMP4, Telechem) was used to print a plasmid DNA/gelatin solution contained in a 384-well plate onto CMT GAPS glass slides (Corning). The pins deposited ~1 nl volumes 400 μm apart using a 25 ms pin down slide time in a 55% relative humidity environment. Printed slides were stored at room temperature in a vacuum desiccator until use. Preparation of aqueous gelatin solution is important and is as follows. 0.02% gelatin (w/v) (Sigma G-9391) was dissolved in MilliQ water by heating and gentle swirling in a 60°C water bath for 15 minutes. The solution was cooled slowly to room temperature and filtered through a 0.45μm cellular acetate membrane and stored at 4°C. Plasmid DNA was purified with the Plasmid Maxi or QIAprep 96 Turbo Miniprep kits (Qiagen), and always had an A260/A280>1.7. Concentrated solutions of DNA were diluted in the gelatin solution so to keep the gelatin concentration >0.017% and, unless otherwise specified, final plasmid DNA concentrations were 0.033 μg/μl. To express GFP the EGFP construct in pBABEpuro was used.

After drying, the DNA spots are briefly exposed to a lipid transfection reagent, the slide is placed in a culture dish and covered with adherent mammalian cells in media. The Effectene transfection kit (301425, Qiagen) was used as follows. In a 1.5 ml microcentrifuge tube, 16 μl enhancer was added to 150 μl EC buffer, mixed, and incubated for 5 minutes at room temperature. 25 μl effectene lipid was added, mixed and the entire volume pipetted onto a 40 x 20 mm cover well (PC200, Grace Bio-Labs). A slide with the printed side down was placed on the cover well such that the solution covers the entire arrayed area while also creating an airtight seal. After a 10 minute incubation, the cover well was pried off the slide with a forceps and the transfection reagent removed carefully by vacuum aspiration. The slide was placed printed side up in a 100 x 100 x 10 mm square tissue culture dish and a 1 x 10^7 actively growing HEK293T cells in 25 ml media (DMEM with 10% FBS, 50 units/ml penicillin and 50 μg/ml streptomycin) were poured into the dish.

Three slides can be transfected side-by-side in this fashion. The cells grew on the slide for 40 hours prior to fixing for 20 minutes at room temperature in 3.7% paraformaldehyde/4.0% sucrose in PBS. Other commonly used mammalian cells lines, such as HeLa and A549 cells, were also tested and similar results were obtained but with transfection efficiencies of 30-50% of those obtained with HEK293 cells. The DNA in the gelatin gel is insoluble in cell culture media but readily enters cells growing on it to create the transfected cell microarray.

To illustrate the method, an array with elements containing an expression construct for the green fluorescent protein (GFP) was printed. HEK293 cells were plated on the slide for transfection and the fluorescence of the cells detected with a laser fluorescence scanner. Microarrays were imaged at a resolution of 5μm with a laser fluorescence scanner (ScanArray 5000, GSI Lumonics). GFP and cy3 emission was measured separately after sequential excitation of the two fluorophores. To obtain images at cellular resolution, cells were photographed with a conventional fluorescent microscope. All images were pseudocolored and superimposed using Photoshop 5.5 (Adobe Systems).

A low magnification scan showed a regular pattern of fluorescent spots that matches the pattern in which the GFP expression construct was printed (Figure 4B). A higher magnification image obtained via fluorescence microscopy showed that each spot is about 150 μm in diameter and consists of a cluster of 30-80 fluorescent cells (Figure 4C). As in a conventional transfection, the total expression level in the clusters is proportional over a defined range to the amount of plasmid DNA used (Figure 4D). Since it may be useful to express two different plasmids in the same cells, whether the technique is compatible with cotransfection was examined. Arrays with elements containing expression constructs for GFP, an epitope-tagged protein or both were prepared and transfected. The cells growing on elements printed with both cDNAs express both encoded proteins, indicating that cotransfection had occurred (Figure 4E).

Whether transfected cell microarrays could be used to clone gene products based on their intrinsic properties was also determined. As a test case, an array to identify the receptor for FK506, a clinically important immunosuppressant whose pharmacologically relevant target, FKBP12, is an intra-cellular protein, was used

(Kino, T., *et al., J. Antiobiot., 40:*1256 (1987); Harding, M.W., *et al., Nature, 26:*755 (1989)). Elements containing expression constructs for FKBP12, GFP, or both were printed on a slide, in an easily recognizable pattern. After the transfected cell microarray formed, radiolabeled FK506 was added to the tissue culture media for one hour prior to processing the slide for autoradiography and immunofluorescence. The radiolabeled FK506 bound to the array in a pattern of spots that exactly matches the pattern of cell clusters expressing FKBP12 (Figure 5A). Detection of the bound FK506 with autoradiographic emulsion confirmed, at the cellular level, colocalization between FKBP12 expression and FK506 binding (Figure 5B). The binding is specific because the GFP-expressing clusters and the non-transfected cells surrounding the clusters showed only background levels of signal (Figure 5A). Furthermore, the prior addition of excess rapamycin, a competitive antagonist of FK506, completely eliminated the signal. 1 μM rapamycin was added to the cell culture media 30 minutes before the addition of radiolabeled FK506.

The utility of transfected cell microarrays for identifying gene products that induce phenotypes of interest in mammalian cells or have a distinct sub-cellular localization was also explored. Arrays with a collection, enriched for signaling molecules, of 192 distinct epitope-tagged cDNAs in expression vectors were printed. 192 Genestrom expression constructs (Invitrogen) in bacteria were cultured in two 96-well plates and plasmid DNA was purified using the Turbo Miniprep Kit (Qiagen). Plasmid DNA was diluted with 0.02% gelatin to a final concentration of 0.040 μg/μl and printed. Cellular phosphotyrosine levels were determined by immunofluorescence staining and scanning. Cell morphology and subcellular localization of expressed proteins was assessed by visual inspection via fluorescence microscopy of the cells in the clusters after their detection with anti-V5 immunofluorescence.

After transfection, their effects on cellular phosphotyrosine levels and morphology as well as their subcellular localization were determined. Five cell clusters on grid 1 (A2, C7, C9, C11, and F6) had phosphotyrosine levels above background (Figure 5C). The coordinates of the clusters match those of the wells of a microtiter plate containing the source cDNAs and were used to look up the identity

of the transfected cDNAs. This revealed that four of these clusters were transfected with known tyrosine kinases (trkC, syk, syn, and blk) while the fifth (C11) encodes a protein of unknown function. Simple visual examination of the morphology of the cells in the transfected clusters revealed a diversity of cellular phenotypes even in this small set of clones. In array 2, cluster E8 had fragmented cells characteristic of apoptosis while in two clusters (D10 and F7) the cells were closely attached to each other (Figure 5D). The presence of apoptotic cells was confirmed by TUNEL (Terminal deoxynucleotidyl transferase mediated dUTP-biotin nick end labeling method) staining. TUNEL staining was performed as described (Y. Gavrieli, Y. Sherman, S.A. Ben-Sasson. J. Cell Biol. 119, 493 (1992)).

The observed phenotypes are consistent with the presumed functions of the cDNAs expressed in these clusters (the Table). Subcellular localization of the expressed proteins were examined through visual inspection the and those with distinct patterns were noted (the Table). This revealed that several proteins that are known transcription factors were mainly located in the cell nucleus. This was also true for other proteins, such as phosphatase 1-beta, whose subcellular distribution has not been previously ascertained.

TABLE

Description of selected cDNAs expressed in the transfected cell microarray. Shown are the coordinates, the phenotype or property detected, the Genbank accession number and the name of the cDNA. nuc/cyto means nuclear and cytoplasmic staining was visible.

Grid: Coordinate	Phenotype/property	Accession number	Function
2:E8	apoptosis	AF016266	TRAIL receptor 2
2:D10	cell adhesion	X97229	NK receptor
2:F7	cell adhesion	M98399	CD36
1:A9	nuclear	U11791	Cyclin H
1:B5	nuclear	M60527	deoxycytidine kinase
1:B12	nuclear	M60724	p70 S6 kinase kinase α1
1:C12	nuclear	M90813	D-type cyclin
1:E4	mitochondrial	U54645	methylmalonyl-coA mutase
1:E10	mitochondrial	J05401	creatine kinase
1:G9	nuc/cyto	U40989	tat interactive protein
1:G10	nuc/cyto	U09578	MAPKAP (3pk) kinase
2:A9	nuclear	X83928	TFIID subunit TAFII28
2:A12	nuc/cyto	M62831	ETR101
2:B6	nuc/cyto	X06948	IgE receptor α-subunit
2:B12	nuclear	X63469	TFIIE β subunit
2:C5	nuclear	M76766	General transcription factor IIB
2:C7	nuc/cyto	M15059	CD23A
2:C12	nuclear	X80910	PP1, β catalytic subunit
2:D4	nuclear	AF017307	Ets-related transcription factor
2:E7	nuclear	X63468	TFIIE α
2:E12	nuclear	U22662	Orphan receptor LXR-α
2:F8	nuclear	L08895	MEF2C
2:F12	nuclear	AF028008	SP1-like transcription factor
2:G2	nuc/cyto	U37352	PP2A, regulatory B' α 1 subunit
2:G3	nuc/cyto	L14778	PP2B, catalytic α subunit

The microarrays can be printed with the same robotic arrayers as traditional DNA arrays, so it is feasible to achieve densities of up 10,000-15,000 cell clusters

per standard slide. At these densities the entire set of human genes can be expressed on a small number of slides, allowing rapid pan-genomic screens. Thus, comprehensive collections of full-length cDNAs for all mammalian genes can be generated (Strausberg, R.L., *et al., Science, 15:*455 (1999); www.hip.harvard.edu/research.html. www.guthrie.org/cDNA.) and will be valuable tools for making such arrays.

Transfected cell microarrays have distinct advantages over conventional expression cloning strategies using FACs or sib selection (Simonsen, H., *et al., Ttrends Pharmacol. Sci., 15:*437 (1994)). First, cDNAs do not need to be isolated from the cells exhibiting the phenotype of interest. This allows for screens using a variety of detection methods, such as autoradiography or *in situ* hybridization, and significantly accelerates the pace of expression cloning. The experiments described herein took days to perform instead of the weeks to months necessary with other expression cloning strategies. Second, transfected cell microarrays can also be used to screen living cells, allowing the detection of transient phenotypes, such as changes in intracellular calcium concentrations. Third, being compact and easy to handle, transfected cell microarrays have economies of scale. The arrays are stable for months and can be printed in large numbers, allowing many phenotypes to be screened in parallel, with a variety of methods, in a small number of tissue culture plates.

Described herein are arrays in which the transfected plasmids direct gene overexpression. However, as antisense technology improves or other methods emerge for decreasing gene function in mammalian cells, it is likely that transfected cell microarrays can be used to screen for phenotypes caused by loss of gene function. Lastly, the immobilization of the plasmid DNA in a degradable gel is the key to spatially restricting transfection without wells.

While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the invention encompassed by the appended claims.

CLAIMS

What is claimed is:

1. A reverse-transfection method of introducing DNA into eukaryotic cells comprising:

(a) depositing a DNA-containing mixture onto a surface in discrete, defined locations, wherein the DNA-containing mixture comprises DNA to be introduced into the eukaryotic cells and a carrier protein and allowing the DNA-containing mixture to dry on the surface, thereby producing a surface having the DNA-containing mixture affixed thereon in discrete, defined locations; and

(b) plating the eukaryotic cells onto the surface in sufficient density and under appropriate conditions for entry of DNA in the DNA-containing mixture into eukaryotic cells,

whereby DNA in the DNA-containing mixture is introduced into the eukaryotic cells.

2. The method of claim 1, wherein the DNA to be introduced is contained in a vector; the carrier protein is gelatin; the slide is a glass slide or a Σ poly-L-lysine slide and the eukaryotic cells are mammalian cells.

3. The method of claim 2, wherein the vector is a plasmid or a viral-based vector.

4. The method of claim 2, wherein the gelatin concentration in the DNA-containing mixture is from about 0.05% to about 0.5%

5. A method of introducing DNA of interest into eukaryotic cells, comprising:

(a) depositing a carrier-DNA mixture onto a surface in discrete, defined locations, wherein the carrier-DNA mixture comprises DNA of interest and a carrier protein, and allowing the carrier-DNA mixture to dry on

the surface, thereby producing a surface bearing the carrier-DNA mixture in discrete defined locations;

(b) covering the surface bearing the carrier-DNA mixture with an appropriate amount of a lipid-based transfection reagent and maintaining the resulting product under conditions appropriate for complex formation between DNA in the carrier-DNA mixture and the transfection reagent;

(c) removing transfection reagent, thereby producing a surface bearing DNA;

(d) plating the eukaryotic cells onto the surface bearing DNA, in sufficient density and under appropriate conditions for entry of the DNA into the eukaryotic cells,

whereby DNA of interest is introduced into the cells.

6. The method of claim 5, wherein the carrier protein is gelatin and the surface is the surface of a slide.

7. The method of claim 6, wherein the slide is a glass slide or a Σ poly-L-lysine slide.

8. The method of claim 7, wherein the concentration of gelatin in the vector-DNA mixture is from about 0.05% to about 0.5%.

9. The method of claim 8, wherein the concentration of gelatin is from about 0.1% to about 0.2%.

10. The method of claim 5, wherein the DNA of interest is in an expression vector and eukaryotic cells that contain DNA of interest are maintained under conditions appropriate for expression of the DNA, whereby DNA of interest is expressed.

11. The method of claim 10, further comprising identifying eukaryotic cells in which a protein of interest is expressed, comprising contacting eukaryotic cells on the surface with an antibody which binds the protein of interest and detecting binding of the antibody, wherein binding identifies eukaryotic cells in which the protein of interest is expressed.

12. A method of introducing DNA of interest into eukaryotic cells, comprising:

(a) depositing a gelatin-DNA mixture onto a surface in discrete, defined locations, wherein the gelatin-DNA mixture comprises DNA of interest and a gelatin, and allowing the gelatin-DNA mixture to dry on the surface, thereby producing a surface bearing the gelatin-DNA mixture in discrete defined locations;

(b) covering the surface bearing the gelatin-DNA mixture with an appropriate amount of a lipid-based transfection reagent and maintaining the resulting product under conditions appropriate for complex formation between DNA in the gelatin-DNA mixture and the transfection reagent;

(c) removing transfection reagent, thereby producing a surface bearing DNA;

(d) plating the eukaryotic cells onto the surface bearing DNA, in sufficient density and under appropriate conditions for entry of the DNA into the eukaryotic cells,

whereby DNA of interest is introduced into the cells.

13. The method of claim 12, wherein the surface is the surface of a slide.

14. The method of claim 13, wherein the slide is a glass slide or a Σ poly-L-lysine slide.

15. The method of claim 14, wherein the concentration of gelatin in the vector-DNA mixture is from about 0.05% to about 0.5%.

16. The method of claim 15, wherein the concentration of gelatin is from about 0.1% to about 0.2%.

17. The method of claim 12, wherein the DNA of interest is in an expression vector and eukaryotic cells that contain DNA of interest are maintained under conditions appropriate for expression of the DNA, whereby DNA of interest is expressed.

18. The method of claim 17, further comprising identifying eukaryotic cells in which a protein of interest is expressed, comprising contacting eukaryotic cells on the surface with an antibody which binds the protein of interest and detecting binding of the antibody, wherein binding identifies eukaryotic cells in which the protein of interest is expressed.

19. The method of claim 4, wherein the eukaryotic cells are mammalian cells and are plated in (b) at high density onto the surface bearing the vector-DNA mixture.

20. A method of introducing DNA of interest into eukaryotic cells, comprising:

(a) depositing a lipid-DNA mixture onto a surface in discrete, defined locations, wherein the lipid-DNA mixture comprises DNA of interest; a carrier protein; a sugar; a buffer that facilitates DNA condensation and an appropriate lipid-based transfection reagent and allowing the lipid-DNA mixture to dry on the surface, thereby producing a surface bearing the lipid-DNA mixture in defined locations;

(b) plating the eukaryotic cells onto the surface bearing the lipid-DNA mixture in sufficient density and under appropriate conditions for entry of DNA of interest into the eukaryotic cells,

whereby DNA of interest is introduced into the cells.

21. The method of claim 20, wherein the carrier protein is gelatin and the surface is the surface of a slide.

22. The method of claim 21, wherein the slide is a glass slide or a Σ poly-L-lysine slide.

23. The method of claim 22, wherein the concentration of gelatin in the lipid-DNA mixture is from about 0.01% to about 0.05% and the concentration of sucrose is from about 0.1M to about 0.4M.

24. The method of claim 20, wherein the DNA of interest is in an expression vector and eukaryotic cells that contain DNA of interest are maintained under conditions appropriate for expression of the DNA, whereby DNA of interest is expressed.

25. A method of affixing DNA to a surface, to produce an array of DNA in discrete, defined locations of known sequence or source, comprising spotting of carrier-DNA mixture onto the surface in discrete, defined locations and allowing the resulting surface bearing the carrier-DNA mixture to dry sufficiently that the spots, referred to as DNA-containing spots, remain affixed to the surface under conditions in which the arrays are used.

26. A method of affixing DNA to a surface, to produce an array of DNA in discrete, defined locations of known sequence or source, comprising spotting of gelatin-DNA mixture onto the surface in discrete, defined locations and allowing the resulting surface bearing the gelatin-DNA mixture to dry sufficiently that the spots, referred to as DNA-containing spots, remain affixed to the surface under conditions in which the arrays are used.

27. A method of affixing DNA to a surface, to produce an array of DNA in discrete, defined locations of known sequence or source, comprising spotting a lipid-DNA mixture onto the surface in discrete, defined locations to produce spots and allowing the resulting surface bearing the lipid-DNA

mixture to dry sufficiently that the spots remain affixed to the surface under conditions in which the arrays are used.

28. A method of producing an array on a surface of reverse transfected cells that contain defined DNA, comprising:

a) spotting a carrier-DNA mixture spotting of gelatin-DNA mixture onto the surface in discrete, defined locations and allowing the resulting surface bearing the carrier-DNA mixture to dry sufficiently that the spots, referred to as DNA-containing spots, remain affixed to the surface under conditions in which the arrays are used;

b) covering the surface bearing the DNA-containing spots with an appropriate amount of a lipid-based transfection reagent and maintaining the resulting product under conditions appropriate for complex formation between DNA in the spots and the transfection reagent;

c) removing transfection reagent, producing a surface bearing DNA;

d) adding cells in an appropriate medium to the surface bearing DNA, to produce a surface bearing DNA and plated cells; and

e) maintaining the surface bearing DNA and plated cells under conditions that result in entry of DNA into plated cells, thus producing an array of reverse transfected cells that contain defined DNA.

29. A method of producing an array on a surface of reverse transfected cells that contain defined DNA, comprising:

a) spotting a gelatin-DNA mixture spotting of gelatin-DNA mixture onto the surface in discrete, defined locations and allowing the resulting surface bearing the gelatin-DNA mixture to dry sufficiently that the spots, referred to as DNA-containing spots, remain affixed to the surface under conditions in which the arrays are used;

b) covering the surface bearing the DNA-containing spots with an appropriate amount of a lipid-based transfection reagent and maintaining the resulting product under conditions appropriate for

complex formation between DNA in the spots and the transfection reagent;

c) removing transfection reagent, producing a surface bearing DNA;

d) adding cells in an appropriate medium to the surface bearing DNA, to produce a surface bearing DNA and plated cells; and

e) maintaining the surface bearing DNA and plated cells under conditions that result in entry of DNA into plated cells, thus producing an array of reverse transfected cells that contain defined DNA.

30. A method of producing on a surface an array of reverse transfected cells that contain defined DNA, comprising:

a) spotting a lipid-DNA mixture onto the surface in discrete, defined locations, to produce spots and allowing the resulting surface bearing the lipid-DNA mixture to dry sufficiently that the spots remain affixed to the surface under conditions in which the arrays are used;

b) plating cells on top of the surface produced in (a) and maintaining the resulting surface, which contains dried lipid-DNA mixture and cells to be reverse transfected, under conditions appropriate for growth of cells and entry of DNA into cells, thus producing an array of reverse transfected cells.

31. An array produced by the method of Claim 25

32. An array produced by the method of Claim 26.

33. An array produced by the method of Claim 27.

34. An array produced by the method of Claim 28.

35. An array produced by the method of Claim 29.

36. An array produced by the method of Claim 30.

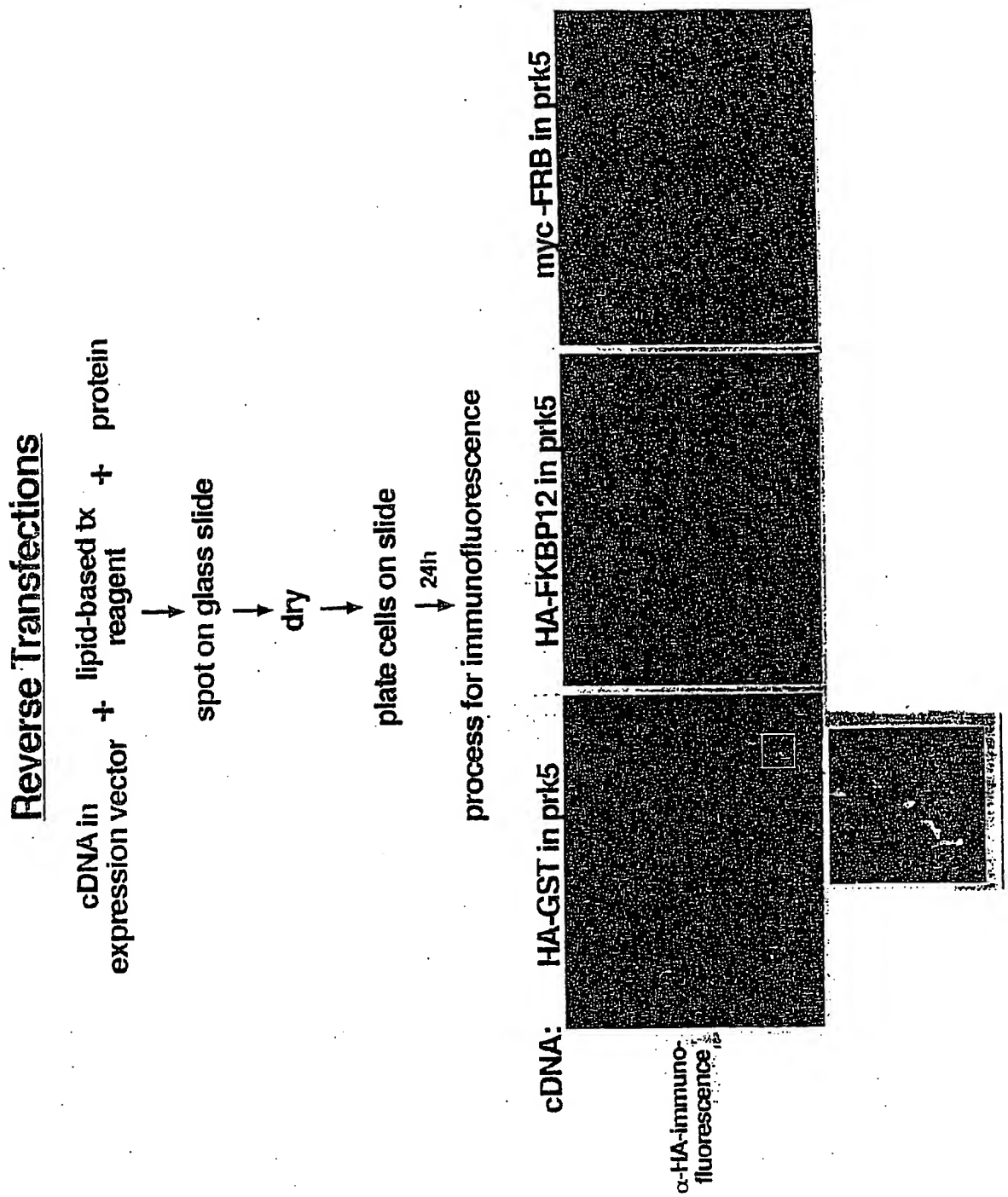

Reverse Transfections
cDNA in expression vector
+
lipid-based tx reagent
+
protein
spot on glass slide
dry
plate cells on slide
24h
process for immunofluorescence
cDNA:
HA-GST in prk5
HA-FKBP12 in prk5
myc-FRB in prk5
α-HA-immuno-fluorescence

Figure 1

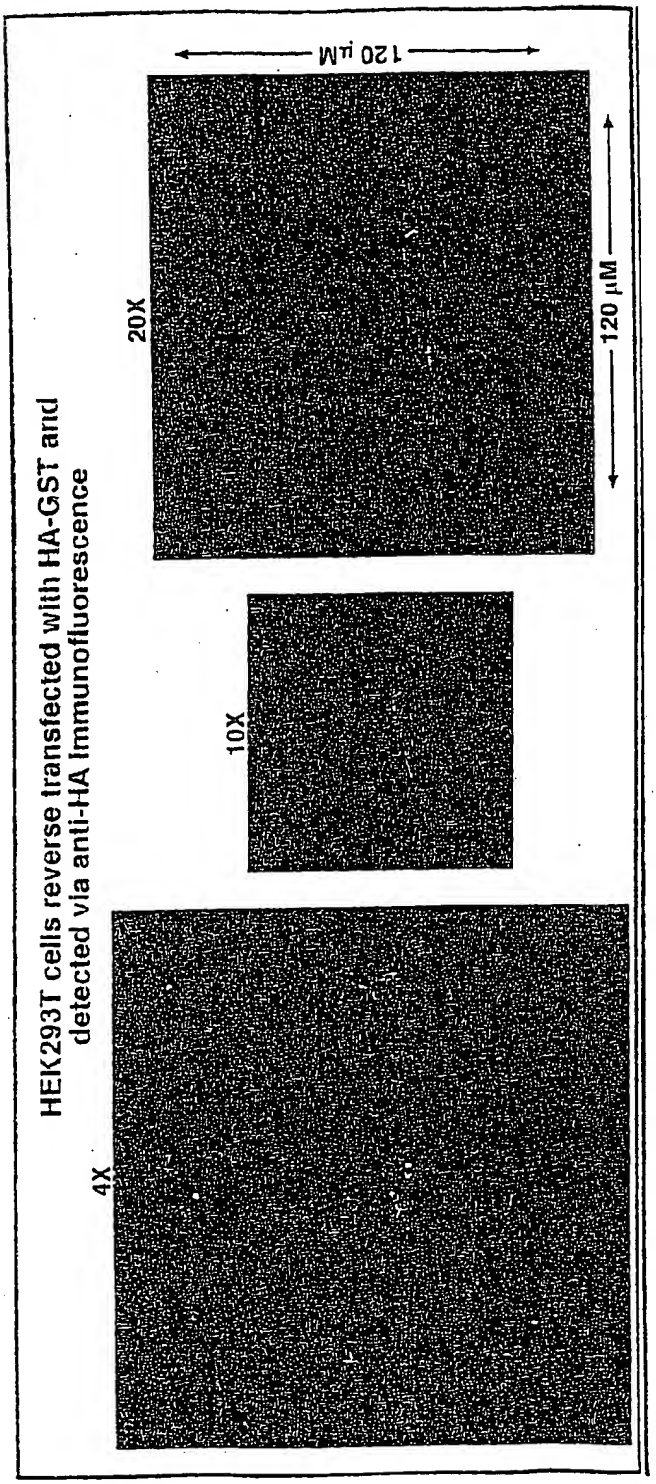

HEK293T cells reverse transfected with HA-GST and detected via anti-HA Immunofluorescence
20X
120 µM
120 µM
10X
4X

Figure 2

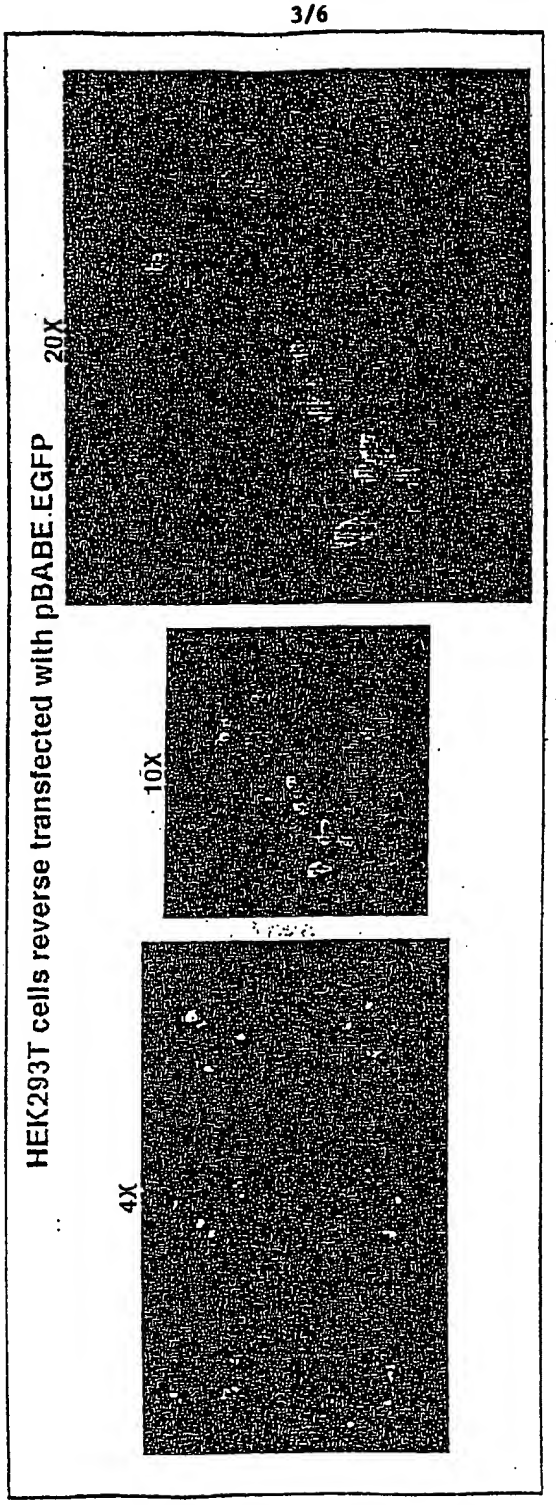

HEK293T cells reverse transfected with pBABE.EGFP
20X
10X
4X

Figure 3

Figure 4 A

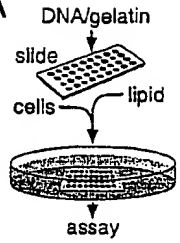


Figure 4 B

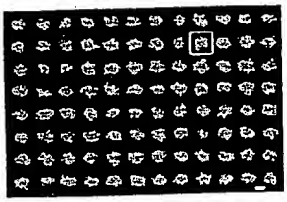


Figure 4 C

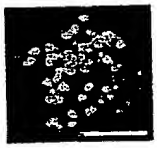


Figure 4D

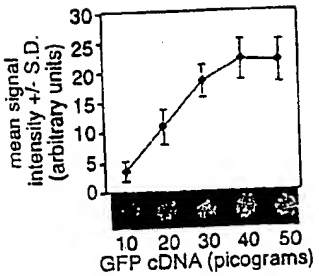


Figure 4E

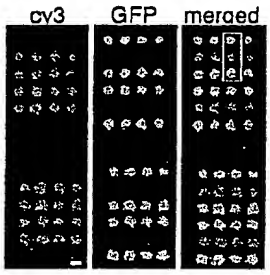


Figure 4F

Figure 5 A

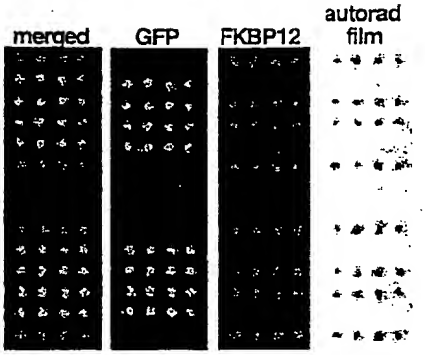


Figure 5 B

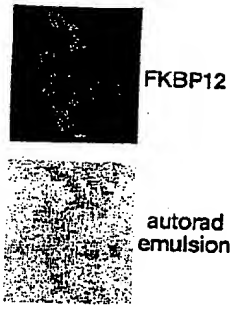


Figure 5 C

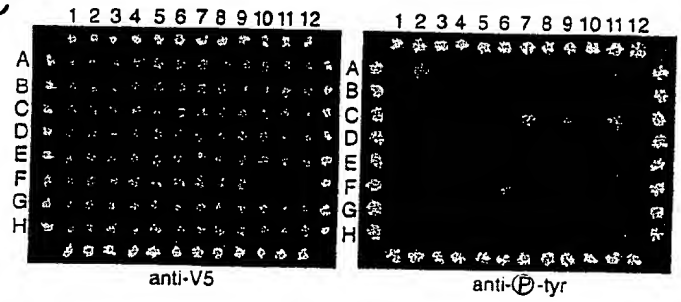
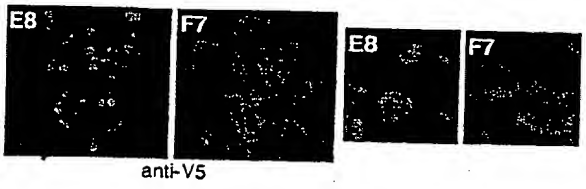


Figure 5 D



INTERNATIONAL SEARCH REPORT

Inten nal Application No
PCT/US 00/25457

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/88 C12N15/10 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 99 55886 A (GENOVA PHARMACEUTICALS CORP) 4 November 1999 (1999-11-04) page 19 -page 20	1
X	WO 95 35505 A (UNIV LELAND STANFORD JUNIOR) 28 December 1995 (1995-12-28) the whole document	25
A	WO 96 17948 A (UNIV MICHIGAN) 13 June 1996 (1996-06-13) page 5, line 29 -page 6, line 5	1-36
A	US 5 851 818 A (OTO EDWIN KIYOSHI ET AL) 22 December 1998 (1998-12-22) abstract column 5, line 28 - line 40	1-36

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 January 2001	17/01/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Galli, I

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter. ...nal Application No
PCT/US 00/25457

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9955886 A	04-11-1999	AU 3572799 A	16-11-1999
WO 9535505 A	28-12-1995	US 5807522 A	15-09-1998
		AT 180570 T	15-06-1999
		AU 709276 B	26-08-1999
		AU 2862995 A	15-01-1996
		CA 2192095 A	28-12-1995
		DE 69509925 D	01-07-1999
		DE 69509925 T	09-12-1999
		DK 804731 T	08-11-1999
		EP 0804731 A	05-11-1997
		EP 0913485 A	06-05-1999
		ES 2134481 T	01-10-1999
		GR 3030430 T	30-09-1999
		JP 10503841 T	07-04-1998
		US 6110426 A	29-08-2000
WO 9617948 A	13-06-1996	US 5811274 A	22-09-1998
		CA 2206600 A	13-06-1996
		EP 0797680 A	01-10-1997
		JP 10510156 T	06-10-1998
		US 5654185 A	05-08-1997
		US 5804431 A	08-09-1998
US 5851818 A	22-12-1998	WO 0015825 A	23-03-2000
		AU 9568698 A	03-04-2000
		US 6133026 A	17-10-2000